

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 15:27:11 ; Search time 31 Seconds
(without alignments)
1180.609 Million cell updates/sec

Title: US-10-055-364-24
Perfect score: 4547
Sequence: 1 MAGSLKRGSLVLAWLWLYQV.....KRSQYSLNKAISEVEMLNTL 865

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents, AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pdp.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pdp.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pdp.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pdp.*
5: /cgn2_6/ptodata/2/iaa/PTCUS_COMB.pdp.*
6: /cgn2_6/ptodata/2/iaa/backfiles.pdp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4547	100.0	865	4	US-09-612-204B-24
2	1953.5	43.0	874	3	US-08-804-439A-16
3	1953.5	43.0	874	3	US-08-720-229-16
4	1949.5	42.9	808	3	US-08-804-439A-14
5	1949.5	42.9	808	3	US-08-720-229-14
6	1937	42.6	845	3	US-08-804-439A-94
7	1937	42.6	845	3	US-08-720-229-94
8	1848.5	40.7	857	3	US-08-220-151-10
9	1848.5	40.7	857	3	US-08-413-118-10
10	1848.5	40.7	857	3	US-08-804-439A-18
11	1848.5	40.7	857	3	US-08-360-107A-113
12	1848.5	40.7	857	3	US-08-473-446-10
13	1848.5	40.7	857	3	US-08-720-229-18
14	1839	40.4	856	3	US-08-486-099-103
15	1839	40.4	856	3	US-08-484-223B-103
16	1839	40.4	856	3	US-08-919-597-103
17	1839	40.4	856	3	US-08-475-668A-103
18	1839	40.4	856	3	US-08-485-551A-103
19	1839	40.4	856	3	US-08-471-913A-103
20	1839	40.4	856	3	US-08-485-264A-103
21	1839	40.4	856	4	US-08-474-349A-103
22	1839	40.4	856	4	US-08-470-896-103
23	1839	40.4	856	4	US-08-485-546A-103
24	1792	39.4	849	3	US-08-804-439A-17
25	1792	39.4	849	3	US-08-720-229-17
26	1787	39.3	874	3	US-08-804-439A-15
27	1787	39.3	874	3	US-08-720-229-15

28	1128	24.8	907	1	US-08-349-006-2	Sequence 2, Appl
29	1128	24.8	907	3	US-08-804-439A-19	Sequence 19, Appl
30	1128	24.8	907	3	US-08-720-229-19	Sequence 19, Appl
31	1128	24.8	907	4	US-09-171-699-2	Sequence 2, Appl
32	1128	24.8	907	5	PCT-US94-04180-2	Sequence 2, Appl
33	1121.5	24.7	906	1	US-08-220-151-9	Sequence 9, Appl
34	1121.5	24.7	906	1	US-08-413-118-9	Sequence 9, Appl
35	1121.5	24.7	906	3	US-08-473-446-9	Sequence 9, Appl
36	1050.5	23.1	830	3	US-08-804-439A-20	Sequence 20, Appl
37	1050.5	23.1	830	3	US-08-720-229-20	Sequence 20, Appl
38	1002	22.0	195	4	US-09-612-204B-2	Sequence 2, Appl
39	901	19.8	879	1	US-08-220-151-2	Sequence 2, Appl
40	901	19.8	879	1	US-08-220-151-3	Sequence 3, Appl
41	901	19.8	879	1	US-08-413-118-3	Sequence 2, Appl
42	901	19.8	879	1	US-08-413-118-3	Sequence 3, Appl
43	901	19.8	879	1	US-08-413-118-106	Sequence 106, App
44	901	19.8	879	3	US-08-473-446-2	Sequence 2, Appl
45	901	19.8	879	3	US-08-473-446-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-09-612-204B-24
; Sequence 24, Application US/09612204B
; Patent No. 6461811
; GENERAL INFORMATION:
; APPLICANT: Patience, Clive
; TITLE OF INVENTION: Swine Gamma Herpesvirus DNA and Methods of Use
; FILE REFERENCE: 61750-299
; CURRENT APPLICATION NUMBER: US/09/612,204B
; CURRENT FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: U.S. 60/142,736
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: U.S. 60/168,532
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 24
; LENGTH: 865
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Deduced amino
; OTHER INFORMATION: acid sequence of porcine gamma herpesvirus gpb
; OTHER INFORMATION: gene
US-09-612-204B-24

Query Match	100.0%;	Score 4547;	DB 4;	Length 865;
Best Local Similarity	100.0%;	Pred. No. 0;	Mismatches 0;	Indels 0;
Matches	865;	Conservative 0;		Gaps 0;
Qy	1	MAGSLKRGSLVLAWLWLYQVALYSLTAETGVTSPPTATWSTESPLTGHYTHDSHGHE	60	
Db	1	MAGSLKRGSLVLAWLWLYQVALYSLTAETGVTSPPTATWSTESPLTGHYTHDSHGHE	60	
Qy	61	RGNNENRDSBEQNNIYGSPSTPPYRVCSAGSGVDVFRFQTDHVCPSADMVHSEGIILLI	120	
Db	61	RGNNENRDSBEQNNIYGSPSTPPYRVCSAGSGVDVFRFQTDHVCPSADMVHSEGIILLI	120	
Qy	121	YKQNIIPFMRVRKRVKVVTTSTVNGIYSDSTINQHTFYKSTIEPTEKMDTIYQCFNS	180	
Db	121	YKQNIIPFMRVRKRVKVVTTSTVNGIYSDSTINQHTFYKSTIEPTEKMDTIYQCFNS	180	
Qy	181	LRLNTGNLLITYVDRDDINMTVFLQPDGVTPDKRYGSOPELYLEPGFWGVSRRRTTV	240	
Db	181	LRLNTGNLLITYVDRDDINMTVFLQPDGVTPDKRYGSOPELYLEPGFWGVSRRRTTV	240	
Qy	241	NCELMDFARSNPFPFFVTTATGDTVEMSPFWSGEDDHENKHEKPMFVSVINNYKVVDY	300	
Db	241	NCELMDFARSNPFPFFVTTATGDTVEMSPFWSGEDDHENKHEKPMFVSVINNYKVVDY	300	

QY 301 QNGRTVPLGKTRIFLDREYYTLTWSKHLKNMSYCPPLTWKAFYNGIQTEHSGSYHFVAND 360
 Db 301 QNGRTVPLGKTRIFLDREYYTLTWSKHLKNMSYCPPLTWKAFYNGIQTEHSGSYHFVAND 360
 QY 361 ITASFTTSKEDMKFNTTYHCLNEEKAEIEKKYAKVNSTHSGYGLKYFKTDGGLYLW 420
 Db 361 ITASFTTSKEDMKFNTTYHCLNEEKAEIEKKYAKVNSTHSGYGLKYFKTDGGLYLW 420
 QY 421 OPLIQNRLLDAKNNLNNETYSRRSQAEISTTDPMMEMTNGAGGYSSENSITVAQVQY 480
 Db 421 OPLIQNRLLDAKNNLNNETYSRRSQAEISTTDPMMEMTNGAGGYSSENSITVAQVQY 480
 QY 481 AYDNLRIIRINILLEDLSKAWCRQHRALVNNLSKINPTSVMSIYNRRVSARKIGDVI 540
 Db 481 AYDNLRIIRINILLEDLSKAWCRQHRALVNNLSKINPTSVMSIYNRRVSARKIGDVI 540
 QY 541 SVSNCIIVDQTSVSLHSLRLLSASDEKCFSRPPVTFKFMNDSTIYKQGLGVNNEILLTT 600
 Db 541 SVSNCIIVDQTSVSLHSLRLLSASDEKCFSRPPVTFKFMNDSTIYKQGLGVNNEILLTT 600
 QY 601 TYLETCOENTYFQAKTDMYIYKNYEHKLTVPGLSSITTLDTFIALNFTLLENVDFKVE 660
 Db 601 TYLETCOENTYFQAKTDMYIYKNYEHKLTVPGLSSITTLDTFIALNFTLLENVDFKVE 660
 QY 661 LYTRDEKRLSNVDFIETMFREYNYAQRVSGLRKDLDDLSTNRNQFVDAFGSLMDDLGA 720
 Db 661 LYTRDEKRLSNVDFIETMFREYNYAQRVSGLRKDLDDLSTNRNQFVDAFGSLMDDLGA 720
 QY 721 GQTVNNAVSGVATLFFSSIVTGFINFKNPFGGMLMIIVIGLVFAIYFLTKTKIYETAP 780
 Db 721 GQTVNNAVSGVATLFFSSIVTGFINFKNPFGGMLMIIVIGLVFAIYFLTKTKIYETAP 780
 QY 781 IKMIYPIDKLKREGKSEIAPISEEELERIVLAMHIHQONSHWEKTRKDPKDSILTRA 840
 Db 781 IKMIYPIDKLKREGKSEIAPISEEELERIVLAMHIHQONSHWEKTRKDPKDSILTRA 840
 QY 841 QNMLRKESGYSLNKNVSEVMLNTL 865
 Db 841 QNMLRKESGYSLNKNVSEVMLNTL 865

RESULT 2

US-08-804-439A-16 Application US/08804439A
 ; Sequence 16, Patent No. 6015565

GENERAL INFORMATION:

; APPLICANT: Rose, Timothy M.
 ; APPLICANT: Bosch, Marnix L.
 ; APPLICANT: Strand, Kurt
 ; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RPHV/KSHV
 ; TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
 ; NUMBER OF SEQUENCES: 113
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson P.C.
 ; STREET: 4225 Executive Square, Ste 1400
 ; CITY: La Jolla
 ; STATE: CA

COUNTRY: USA

ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

CURRENT APPLICATION DATA:

SOFTWARE: PatentIn Release #1.0, Version #1.30

APPLICATION NUMBER: US/08/804.439A

FILING DATE: February 21, 1997

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Haile, Lisa A.

REGISTRATION NUMBER: 38,347

REFERENCE/DOCKET NUMBER: 09176/004001

TELECOMMUNICATION INFORMATION:

; TELEPHONE: (619) 678-5070
 ; TELEFAX: (619) 678-5099
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 16:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 874 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-804-439A-16

Query Match 43.0%; Score 1953.5; DB 3; Length 874;

Best Local Similarity 46.1%; Pred. No. 2e-143;

Matches 403; Conservative 147; Mismatches 250; Indels 75; Gaps 18;

QY 11 VLALWLYQVALYSLSIAETGYTSPENTATWSTESPLTGHYGFHDSHGGRGNENRDS 70
 Db 10 VLCLWCV--AALLCQGAQEVVA-----ETTTB---FATH-----RPEVVAE 46
 QY 71 EQKNIIYGSPTFPYRVCASGV-GDVPRFQDHYCPDASDMVHSEGIILLIYKQNIIPFM 129
 Db 47 ENPAMPF-----LPRFVCGASPTGGEIFRPLEESCPTEDKDHIEGIALIYKNTIYVPV 101
 QY 130 FRVRKYRKVVTTSTVYNGIYSDSITNQHTFYKSIPEWETEKMDTIYQCNSLRNLNTGSL 189
 Db 102 FNVRYKRYKMTSTIYKHSSEDAITNQHTRSYAVPLYEYVQMDHYHYQCFSAVQVNEGSHV 161
 QY 190 LTYVDRDDINMTVFFQPDVGVTPOVKRYGSGPELYLEPCGMFGWSYRRRTTIVNCELMDFA 249
 Db 162 NTYVDRDGNWETAFLKPADGLTSSITRYQSQPEVYATPRNLLWSYTRITVNCVEVTEMSA 221
 QY 250 RSNPPDFPVATGDTVENSFPWSGEDDHENKMKHEKWPVSVINNYKVVDYQNRGTVPGLG 309
 Db 222 RSMKPEPFVTSVGDGTIENSFPFLKENGTEPEKILRPHSIIQLLKNVAVTKY-----GVGLG 277
 QY 310 K---TRIFLDREYTLWSKHLKNMSYCPLTLWKAFYNGIQTEHSGSYHFVANDITASF 365
 Db 278 QADNATRFPAIFGDSLSLWKATENSYYCDLLWKGFSAIQTQHNSSLHFANDITASF 337
 QY 366 TTSKEDMKFNTTYHCLNEEKAEIEKKYAKVNSTHSGYGLKYFKTDGGLYLWVQPLIQ 425
 Db 338 STPLEBEANFNFTFKCIWNNTQBEIKLKEVEKTHRPNGTAKVYKTTGNLYIVWQPLIQ 397
 QY 426 NRLLDAKNNLNN-----ETYSRRSRQAEISTTDPMMEMTNGAGGYSSENS 472
 Db 398 IDLLDTHAKLYNLNTATASPTSTPTSPRRRRDTS-----VSGGNGNGDNSTKEES 450
 QY 473 ITVAQVOYAYDNLRIIRINILLEDLSKAWCRQHRALVNNLSKINPTSVMSIYNRPVS 532
 Db 451 VAASQVQFAYDNLKSIINRVLGELSRAWCREQYRASLMWYELSKINPTSVMSIAYGEPVS 510
 QY 533 AKRIGDVI SVSNCIIVDQTSVSLHSLRLLSASDEKCFSRPPVTFKFMNDSTIYKQGLGV 592
 Db 511 AKLIGDVSVSDCISVDQSVFVHKNMK-VPGKEDLCYTRPVVGVFKFNGSELFPAGQLGP 569
 QY 593 NNEILLTITYTCQENTYFQAKTDMYIYKNYEHKLTVPGLSSITTLDTFIALNFTLLE 652
 Db 570 RNEIVLSTSQVEYCOHSCHEHYFOAGNQMKYKDYVYSTNLNLTDTPLTHTMITNLNLSVE 629
 QY 653 NYDFKVIELYTRDEKRLSNVDFIETMFREYNYAQRVSGLRKDLDDLSTNRNQFVDAFG 711
 Db 630 NIDFKVIELYKTEKRLSNVDFIETMFREYNYQNLNGLRKDLDDLSDIDHGRSFTQTLG 689
 QY 712 SLMDLLGAVGQTVNNAVSGVATLFFSSIVTGFINFKNPFGGMLMIIVIGLVFAIYFLT 770
 Db 690 DIMQDLGTIGKVVNVNAVSGVSLFGSIVSGVISFFKNPFGGMLLIIVLIIAGVVVYLFMT 749
 QY 771 KTKIYETAPIKMIYEPIDKLKERECKSEIAPISEEELERIVLAMHIHQONSHWEKTRK 830
 Db 750 RRSIY-SAPIRMLYPGVERAAQEPGAH---PVSEDIIRNLMGMHGFQQRARAEBAER 805
 QY 831 DPK-----DSILTRAQNMLRKR---SGYSNLK 854


```

; TELEFAX: (619) 678-5099
; TELEX:
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 808 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
; US-08-804-439A-14

Query Match          42.9%; Score 1949.5; DB 3; Length 808;
Best Local Similarity 46.6%; Pred. No. 3.6e-143;
Matches 373; Conservative 159; Mismatches 236; Indels 33; Gaps 10;

QY 71 EQNKNIYGSSTPYRVCSASGVGVDFRQTDHVCPCDASDMVHSEGLLIYKQNIIPFPMF 130
Db 28 EKNKTOAIYQEFYKRVCSASITGELFRDLRTPSTEDKVHKEGILLVYKKNIVPIYIF 87
QY 131 RVKRYKVVTTSTVYNGIYSD--SITNOHTFYKSIPEWETEKMDTIYQCFNSRLNTGN 188
Db 88 KVRKYKAITTSVRIFNGWTRREGVAITNKWELSRAPVKYEIDIMDKTYQCHNCQIEVNGM 147
QY 189 LLTYVDRDDINMTVLPQVGVTPDVKRYGSQPELYLEPGWFGSVRRRTTVNCELMDMF 248
Db 148 LNSYDGRGNKTVLDKPVDTGUTGAIIRYISQPKVADPGWLWGTTRTTTVNCEIVDMF 207
QY 249 ARSNPPDFVFTATGDTVEMSPFSGEDDHNKMEKPMFVSVINNYKVVDYQNRGTVPL 308
Db 208 ARSADPYTYFTALGDTVEVSPFCDVDNSCPNATDVL--VQIDLNHTVVDYGNRATSOQ 265
QY 309 GKTRIFLDREYTLNWEKHLKNMSYCLPLTWKAFYNGIOTHSYGVHVFANDITASFTTS 368
Db 266 HKKRIFAHTLDYSVSWAEAVNKASVCMSVFWKSFQRAIQTEHDLTVHFANEITAGFSTV 325
QY 369 KEDMKEFNNTYHCLNEEIKAEIEKKYAKVNSTHSGYDLKYEKTDGGLVLMQPLIQNRL 428
Db 326 KEPLANFTSDYNCLMTHINTTLEDKIARVNNHTPNGTAEYQTEGGMLLVWQPLIAIEL 385
QY 429 LDANKLANNE-----TYSRRSRROAESTTDPMMETNGAGGEYSSENSITVAQVOYA 481
Db 386 EAMLEATTSPVTPSAPTSSRSKRRAIRSIRDV-----SAG-----SENNVFLSQIOYA 434
QY 482 YDNLRIINLLEDSKAWCREQHRALVNELSKINPTSVMSMIYNRPVSAKRIGDVIS 541
Db 435 YDKLRQSIINNVLEBALTWCREQVRQTVWVYEAIAKINPTSVMTAIVGKPVSRKALGDVIS 494
QY 542 VSMCIVVDQTSVSLHKSRLLSASDEKCFRPPVTFKPMNDSTIYKGQLGVNNEILLTTT 601
Db 495 VTECINVDQSSVSIHSLK--TENNDICYSRPPVTFKFNSSQLFKGQLGARNEILLSES 552
QY 602 YLETCQENTYYPQAKTDMYIKNYEHLKTVPLSSITLDTFLALNPTLLENVDFKVEL 661
Db 553 LVENCHQNAETFTAKNETYHFKNYVHVTLPVNNISTLDTFLALNLTFTENIDFKAVEL 612
QY 662 YTRDEKSLNVFIDETMFRYNYVYQAVSGLRDLDLDS--TNRNQVDAFGSLMDDLGA 719
Db 613 YSSGERKANVFDLETMFRYNYVYQAVSGLRDK--FDNSQNRNDRDIIOQFSEILDGLS 671
QY 720 VQOTVNVAVSGVATLFSFIVTGFINFKNPFGGMLMIIVVIGLVFAIYFLTKTKIYETA 779
Db 672 IGKVINNVASGAFSLFGGIVTGILNFKNPLGGMFTFLIGAVILLVILLVRRTNNMSQA 731
QY 780 PIKMIYPIEKLKEREKKSIAPISEBELRIVLAMIHQONSHMETKTRKDPKDSILTR 839
Db 732 PIRMIYDPVFK-----SKSTVTWPEPETIKQILLGMHMQOEAQYKKKEQRAARPSIFRQ 786
QY 840 -AQNMLKRSYSNKLNAESV 859
Db 787 AAEFTLKRSGYKQISTEDKI 807

```

```

US-08-720-229-14
; Sequence 14, Application US/08720229
; Patent No. 6022542
; GENERAL INFORMATION:
;   APPLICANT: Rose, Timothy M.
;   APPLICANT: Bosch, Marnix L.
;   APPLICANT: Strand, Kurt
;   TITLE OF INVENTION: GLYCOPROTEIN B OF THE RPHV/KSHV
;   TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
;   NUMBER OF SEQUENCES: 100
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Morrison & Foerster
;     STREET: 755 Page Mill Road
;     CITY: Palo Alto
;     STATE: CA
;     COUNTRY: USA
;     ZIP: 94304-1018
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: Patentin Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/720,229
;     FILING DATE: 26-SEP-1996
;     CLASSIFICATION: 424
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Schiff, J. Michael
;     REGISTRATION NUMBER: 40,253
;     REFERENCE/DOCKET NUMBER: 29938-20002.00
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (415) 813-5600
;     TELEFAX: (415) 494-0792
;     TELEX: 706141
;   INFORMATION FOR SEQ ID NO: 14:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 808 amino acids
;       TYPE: amino acid
;       STRANDEDNESS: single
;       TOPOLOGY: linear
;       MOLECULE TYPE: protein
;   US-08-720-229-14

```

```

Query Match          42.9%; Score 1949.5; DB 3; Length 808;
Best Local Similarity 46.6%; Pred. No. 3.6e-143;
Matches 373; Conservative 159; Mismatches 236; Indels 33; Gaps 10;

QY 71 EQNKNIYGSSTPYRVCSASGVGVDFRQTDHVCPCDASDMVHSEGLLIYKQNIIPFPMF 130
Db 28 EKNKTOAIYQEFYKRVCSASITGELFRDLRTPSTEDKVHKEGILLVYKKNIVPIYIF 87
QY 131 RVKRYKVVTTSTVYNGIYSD--SITNOHTFYKSIPEWETEKMDTIYQCFNSRLNTGN 188
Db 88 KVRKYKAITTSVRIFNGWTRREGVAITNKWELSRAPVKYEIDIMDKTYQCHNCQIEVNGM 147
QY 189 LLTYVDRDDINMTVLPQVGVTPDVKRYGSQPELYLEPGWFGSVRRRTTVNCELMDMF 248
Db 148 LNSYDGRGNKTVLDKPVDTGUTGAIIRYISQPKVADPGWLWGTTRTTTVNCEIVDMF 207
QY 249 ARSNPPDFVFTATGDTVEMSPFSGEDDHNKMEKPMFVSVINNYKVVDYQNRGTVPL 308
Db 208 ARSADPYTYFTALGDTVEVSPFCDVDNSCPNATDVL--VQIDLNHTVVDYGNRATSOQ 265
QY 309 GKTRIFLDREYTLNWEKHLKNMSYCLPLTWKAFYNGIOTHSYGVHVFANDITASFTTS 368
Db 266 HKKRIFAHTLDYSVSWAEAVNKASVCMSVFWKSFQRAIQTEHDLTVHFANEITAGFSTV 325
QY 369 KEDMKEFNNTYHCLNEEIKAEIEKKYAKVNSTHSGYDLKYEKTDGGLVLMQPLIQNRL 428
Db 326 KEPLANFTSDYNCLMTHINTTLEDKIARVNNHTPNGTAEYQTEGGMLLVWQPLIAIEL 385
QY 429 LDANKLANNE-----TYSRRSRROAESTTDPMMETNGAGGEYSSENSITVAQVOYA 481

```


Db 386 BEAMLEATTPTVPAPTSSSKRAIRSIRDV-----SAG-----SENNVFLSQIOYA 434
Qy 482 YNLRIRINNIILEDISKACREQRAALVWNLKINPTSMVSMIYNRPVSAKRIGDVIS 541
Db 435 YDKLRQSNINVLLELAITWCRCQVQRTWVYIAKINPTSMVIAIYGRPVSKALGDVIS 494
Qy 542 VENCIVVDQTSVSLHKSRLLSASDEKCFRPPVTFKFMNDSTIYKGLGVNNEILLTTT 601
Db 495 VTECINVDQSSVSIHSLK--TENNDICVSRPPVTFKFNSSQLFKGQLGARNEILLSES 552
Qy 602 YLETCOENETYYFOAKTDMYIYKNVEHLKTVPLSSITLDTFIALNFTLLENVDFKVEL 661
Db 553 LVENCHQNAETFTAKNETHPKNVHVETLEPVNNISLDTFLALNLFTIENIDFKAVEL 612
Qy 662 YTRDEKRLSNVFDIETMPEYNNYQVRVSGRLKLDLS--FNROQVDAFGSLMDLGA 719
Db 613 YSSGERKANVDELTMPEYNNYQVSIQSLKRD-FDMSQRNRRRIIQDFSEILLADJGS 671
Qy 720 VQGTVVNAVSGVATLFSSTIVTGFINFKNPFQGLMLMIIVIGVLFAIYFLTKTKIYETA 779
Db 672 IGKVIWNVASGAFSLFGGIVTGILNFKNPLGGMFTFLIGAVIILVILLVRRTNMMSQA 731
Qy 780 PIKMIYPELDKLEREGKSEIAPISEEELERIVLAHMHQONSHMETKTRKDPKDSILTR 839
Db 732 PIRMIYDVEK-----SKSTVTMPPEPETIKQILLGMHNMQOEAYKKKEEORAAARPSIFRQ 786
Qy 840 -AQNMLRKRSGYSLKNAESV 859
Db 787 AAETFLRKRSGYKQISTEDKI 807

RESULT 6

US-08-804-439A-94
; Sequence 94, Application US/08804439A
; Patent No. 6015565
; GENERAL INFORMATION:
; APPLICANT: Rose, Timothy M.
; APPLICANT: Bosch, Marnix L.
; APPLICANT: Strand, Kurt
; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RPHV/KSHV
; TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Ste 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,439A
; FILING DATE: February 21, 1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 09176/004001
; TELEPHONE: (619) 678-5070
; TELEFAX: (619) 678-5099
; TELEX:
; INFORMATION FOR SEQ ID NO: 94:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 845 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:

; NAME/KEY: Modified-site
; LOCATION: 841
; OTHER INFORMATION: /note= "Proline or Leucine
; OTHER INFORMATION: depending on codon"
US-08-804-439A-94

Query Match 42.6%; Score 1937; DB 3; Length 845;
Best Local Similarity 48.5%; Pred. No. 3.7e-142;
Matches 382; Conservative 135; Mismatches 237; Indels 34; Gaps 10;

Qy 83 PPRVCSASGVDPVFRQTDRHVCDDASDMVHSEGLLIYKONIIPFPRVRKYKRVVTS 142
Db 63 YQFVCSASITGELFRFNLEQTCPTDKVKHQEGILLVYKKNIVPHIFKVRRYKRIATSV 122
Qy 143 TVYNGIYSDSTTNOHTFYKSIPEWETEKMDTIYOCFNSLRNLNTGNTLLTYVDRODINMTV 202
Db 123 TVYRGLTESAITNKEYELPRPVPLYEISHMDSTYQCFSSMKVNVNGVENTFTFDRDDVNTTV 182
Qy 203 FLQPVGVTPDVKRYGSOPELYLEPFGWFSYRRRTTVNCELMDMFARNSNPPFFVVTAT 262
Db 183 FLQPVEGLTDNIQRYFQPVYAEFGWFPGLYRVRTTVNCEIVDMIARSAEPYNYFVTS 242
Qy 263 GDTVEMSPFSGDEDDHKNMHEKMPF-VSVINNVYVDYQNRGTVPGLKTRIFLDREY 321
Db 243 GDTVEVSPFCYNSSCSTTPSNKNGLSVQVNLNHTVTVYSDRGTSPQNRIFVETGAYT 302
Qy 322 LSWEKHLKNMSYCPILTAKAFYNGIOTHSYHYFVANDITASFTTSKEDMKFNTTVC 381
Db 303 LSWASESKTTAVCPALWKTTPRSIQTHEDSPHFVANEITATFTAPLTPVANFTDVTSC 362
Qy 382 LNEEIKAEIEKKYAKVNSTHSKYGDLKYFKTGDGLYLVWQPLIQNRLLDAKNKLNNET-- 439
Db 363 LQSDINTLNASKAKLASTHVPNGTVQYFHTTGLYLVWQPMASINLTHAGDSGNPTSS 422
Qy 440 -----YSRRSRQAESTTDPMMWMTGNAGGEGYSENSITVAQVQYADNLRIR 488
Db 423 PPSASPMWTTASRRKRSSASTA-----AAGGGSTDN-LSVTQLQFAYDKLRDG 471
Qy 489 INNILEDLSKAWCEQRAALVWNLKINPTSMVSMIYNRPVSAKRIGDVISVNCIVV 548
Db 472 INQVLEELSRACREQVRDNLMMYELSKINPTSMVIAIYGRPVSAKFVGDAISVTECINV 531
Qy 549 DQTSVSLHKSRLLSASDEKCFRPPVTFKFMNDSTIYKGLGVNNEILLTTTLETCE 608
Db 532 DQSSVNIHKSRL--TNSKDVYARPLVTFKLNSSNLTGOLGARNELIILTNNOVEICKD 589
Qy 609 NTEYFOAKTDMYIYKNVEHLKTVPLSSITLDTFIALNFTLLENVDFKVELYTRDEKR 668
Db 590 TCEHYFITRNETLVYKDYAVLRTINTTDTLNTFIALNLSFIQNIIDFKALIELYSSAEKR 649
Qy 669 L-SNVFDIETMPEYNNYQVRVSGRLKLDL--LSTNRNQVDAFGSLMDLGAQVQTVN 726
Db 650 LASSVFDLETMPEYNNYTHRLAGLRDLDTDMNKRFRVLDSEIVADLGGIGKTVN 709
Qy 727 AVSGVATLFSSTIVTGFINFKNPFQGLMLMIIVIGVLFAIYFLTKTKIYETAPIKMIY 786
Db 710 VASSVTLGSLVTCGFINFKPLGGLMLMIIVIAIILIFMLSRRTNTTAQAPVKMIY 769
Qy 787 EIDKLKEREGKSEIAPISEEELERIVLAHMHQONSHMETKTRKDPKDSILTRQANMLRK 846
Db 770 DVDR--RAPPSSGAP-TREEIKNILLGMHOLQOEERQKADDLKSKTSPSFQRTANGLRQ 825
Qy 847 R-SQYSNL 853
Db 826 RLRGYKPL 833

RESULT 7

US-08-720-229-94
; Sequence 94, Application US/08720229
; Patent No. 6022542
; GENERAL INFORMATION:
; APPLICANT: Rose, Timothy M.


```
Query Match      40.7%; Score 1848.5; DB 1; Length 857;
Best Local Similarity 46.1%; Pred. No. 2.9e-135;
Matches 368; Conservative 148; Mismatches 241; Indels 41; Gaps 12;

Qy 81 STFPYRVCSAGVGDFRQTDHVCDD-ASDMVHSEGLLIYKQNIIPFMRVRKRVK 139
Db 44 TSFPRVCELSSHGLDFRSDIQCPSTRENHTEGLLMVFKDNIIPYSFKVRSYTKIV 103
Qy 140 TTSTVYNGIYSDSTNQHTEYKSIPEWETEKMDTIYOCNSLRNLNTGNNLLTYVDRDIN 199
Db 104 TNILYNGWADSVNRHEEKFSDSYETDQMDTIYOCYNAVMTKDGLTRVYVDRDGVN 163
Qy 200 MTFVLQPDVGVTPDKRYGQPELYLEPGWFGSYRRRTTNCLEMDMFARNSPPDFV 259
Db 164 ITVNLKPTGGLANGVRVYASQTELYDAPGWLITWYRTTIVNCLITDMAKSNPFFDFV 223
Qy 260 TATGDTVEMSPFSGEDDHENKMKHKEPWFVSINNYKVVDYQNRGTVPGLKTRIFLDREE 319
Db 224 TTTGQTVEMSPFYDGK--NKETHERADSFHVRTNYKIVDYNRGTNPQGERAFDLKGT 281
Qy 320 YTLSEKHLKN-MSYCPILTAKAFYNGIOTESHGSHYFVANDITASFTTSKEDMKFNTT 378
Db 282 YTLSEKHLKN-MSYCPILTAKAFYNGIOTESHGSHYFVANDITASFTTSKEDMKFNTT 378
Qy 379 YHCLNEEIKAEIEKKYAKVNSTHSHKYGD-LKYFKTDGGLYLVWQPLIQNRLDKN--KL 435
Db 340 FKIEEQVNTKHEKYEAVQDRTYKQEAITYFITSGLLLAWLPLTPRSLATVKNLTEL 399
Qy 436 NNETY-----RRSRQAESTTDPMMEMTNGAGGESSNSI 473
Db 400 TPTSPSPSPSPAPSAARGSTPAALVLRRRRRDAGNATP---VPPTAPGKSLGLTNNP 456
Qy 474 TVAQVQYAYDNLIRINNILEDLKAWCREQRAALVWNLKINPTSVMSIYNRPVSA 533
Db 457 ATVQIQFAYDSLRRQINRMGLDLARAWCLEQKQNMVRLTKINPTTVMSSYIGKAVAA 516
Qy 534 KRIGDIVSNCIVVDQTSVSLHSLRLLSASDEKCFSPRPVTPKFMNDSTIYKGLGVN 593
Db 517 KRLGDIVSNCIVVDQTSVSLHSLRLLSASDEKCFSPRPVTPKFMNDSTIYKGLGVN 593
Qy 594 NEILLTTLTYLETCQNTYFYQAKTDMYIKYNEHLKTVPLSSITLDTFIALNFTLLEN 653
Db 576 NEIFLTQWTEVCQATSOYFQSGNEIHVYNDYHFKTIELDGIATLQTFISLNTSLIEN 635
Qy 654 VDFKVIELYTDEKRLSNVFIETMFRYNYAQVSLGRKDLDD-LSTNRNQFVDPAGS 712
Db 636 IDFASLELYSDEQASNVDFLEGIFREYNFQAQNIAGLRKDLONAVSNGRNQFVGLGE 695
Qy 713 LMDLLGAGQTVNAVSGVATLFFSSIVTGFINFKNPFGLMIIIVIGLVFAIYFLTKK 772
Db 696 LMDLSGSGVQITNLVSTVGGFLSLSGFIISFFKNPFGGLMIIIVAGVILVISLRR 755
Qy 773 TKIVETAPIKMYPEIDKLKEREKSE---IAPISEEELERIVLAMHIQOONSHMETKR 829
Db 756 TRQMSQOQVOMLYPGIDELAQHASGEGGINPISKTELQAIMLA--LHEQNOQKRAAQ 813
Qy 810 KDPKDSILTRAQNMRLKR 847
Db 814 RAAGPSVASRALQAARD 831
```

RESULT 9

```
US-08-413-118-10
; Sequence 10, Application US/08413118
; Patent No. 5688920
; GENERAL INFORMATION:
; APPLICANT: PAOLETTI, ENZO
; APPLICANT: LIMBACH, KEITH J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; CANINE HERPESVIRUS gb, gc, AND gd AND USES THEREFOR
; NUMBER OF SEQUENCES: 128
; CORRESPONDENCE ADDRESS:
```

```
; ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
; STREET: 530 FIFTH AVENUE, 25TH FLOOR
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/413,118
; FILING DATE: 29-MAR-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/220,151
; FILING DATE: 30-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: FROMMER, WILLIAM S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2670
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 857 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
; US-08-413-118-10
```

```
Query Match      40.7%; Score 1848.5; DB 1; Length 857;
Best Local Similarity 46.1%; Pred. No. 2.9e-135;
Matches 368; Conservative 148; Mismatches 241; Indels 41; Gaps 12;

Qy 81 STFPYRVCSAGVGDFRQTDHVCDD-ASDMVHSEGLLIYKQNIIPFMRVRKRVK 139
Db 44 TSFPRVCELSSHGLDFRSDIQCPSTRENHTEGLLMVFKDNIIPYSFKVRSYTKIV 103
Qy 140 TTSTVYNGIYSDSTNQHTEYKSIPEWETEKMDTIYOCNSLRNLNTGNNLLTYVDRDIN 199
Db 104 TNILYNGWADSVNRHEEKFSDSYETDQMDTIYOCYNAVMTKDGLTRVYVDRDGVN 163
Qy 200 MTFVLQPDVGVTPDKRYGQPELYLEPGWFGSYRRRTTNCLEMDMFARNSPPDFV 259
Db 164 ITVNLKPTGGLANGVRVYASQTELYDAPGWLITWYRTTIVNCLITDMAKSNPFFDFV 223
Qy 260 TATGDTVEMSPFSGEDDHENKMKHKEPWFVSINNYKVVDYQNRGTVPGLKTRIFLDREE 319
Db 224 TTTGQTVEMSPFYDGK--NKETHERADSFHVRTNYKIVDYNRGTNPQGERAFDLKGT 281
Qy 320 YTLSEKHLKN-MSYCPILTAKAFYNGIOTESHGSHYFVANDITASFTTSKEDMKFNTT 378
Db 282 YTLSEKHLKN-MSYCPILTAKAFYNGIOTESHGSHYFVANDITASFTTSKEDMKFNTT 378
Qy 379 YHCLNEEIKAEIEKKYAKVNSTHSHKYGD-LKYFKTDGGLYLVWQPLIQNRLDKN--KL 435
Db 340 FKIEEQVNTKHEKYEAVQDRTYKQEAITYFITSGLLLAWLPLTPRSLATVKNLTEL 399
Qy 436 NNETY-----RRSRQAESTTDPMMEMTNGAGGESSNSI 473
Db 400 TPTSPSPSPSPAPSAARGSTPAALVLRRRRRDAGNATP---VPPTAPGKSLGLTNNP 456
Qy 474 TVAQVQYAYDNLIRINNILEDLKAWCREQRAALVWNLKINPTSVMSIYNRPVSA 533
Db 457 ATVQIQFAYDSLRRQINRMGLDLARAWCLEQKQNMVRLTKINPTTVMSSYIGKAVAA 516
Qy 534 KRIGDIVSNCIVVDQTSVSLHSLRLLSASDEKCFSPRPVTPKFMNDSTIYKGLGVN 593
```


	STREET:	1155 Avenue of the Americas		
	CITY:	New York		
	STATE:	New York		
	COUNTRY:	USA		
	ZIP:	10036-2711		
	COMPUTER READABLE FORM:			
	MEDIUM TYPE:	Floppy disk		
	OPERATING SYSTEM:	IBM PC compatible		
	SOFTWARE:	Patent In Release #1.0, Version #1.30		
	CURRENT APPLICATION DATA:			
	APPLICATION NUMBER:	US/08/360,107A		
	FILING DATE:	20-DEC-1994		
	CLASSIFICATION:	435		
	ATTORNEY/AGENT INFORMATION:			
	NAME:	Coruzzi, Laura A.		
	REGISTRATION NUMBER:	30,742		
	REFERENCE/DOCKET NUMBER:	7872-013		
	TELECOMMUNICATION INFORMATION:			
	TELEPHONE:	(212) 790-9090		
	TELEX:	66141 PENNIE		
	INFORMATION FOR SEQ ID NO:	113:		
	SEQUENCE CHARACTERISTICS:			
	LENGTH:	857 amino acids		
	TYPE:	amino acid		
	STRANDEDNESS:	unknown		
	TOPOLOGY:	protein		
	MOLECULE TYPE:	protein		
	US-08-360-107A-113			
	Query Match	40.7%; Score 1848.5; DB 3;	Length 857;	
	Best Local Similarity	46.1%; Pred. No. 2.9e-135;		
	Matches	368; Conservative 148; Mismatches 241;	Indels 41; Gaps 12;	
Qy	81	STFPRVCSASGVDFRQTTHVCDD--ASDWMHSEGLLIYKQNIIPMFRVRKYRKVV	139	
Dd	44	TSPFFRVCELSSGDLRFESSDIQCPSFGTRENHTEGLLMVFKDNIIPIYSFKRSYTIV	103	
Qy	140	TTSTVNNGIYSDITNOHTFYKSIBETWEKMDTIYCENSURLNTGNLLTVDRDDIN	199	
Dd	104	TNILLINGVADSVTRHEEKFVSVDYEYDMDTIYCYNAMVKDKGLTRYVVDROGVN	163	
Qy	200	MTVLFPDVGGVTDPVKRYGSQPLEYLPDGWFMGSRRTTNVCELMDFARSNPPFDFFV	259	
Dd	164	ITVNLKPTGSLANGRRVASQTELYPDAGCWLWTYRTRITVNCILTDMMAKSNSPFDFFV	223	
Qy	260	TATGDTVENSGPWGGEDHENKMKEKPWFVSVNNYKVVDYQNQRGTVPVGKTRIPLDREE	319	
Dd	224	TTTTGTVMESPFYDGR--NKETFHERADSFHVRTNYKIYDYDNRGNTNQGERAPLDKGT	281	
Qy	320	YTLWSKKHLKN-MSYCLPLTKAFYNGIOETHSGSHVFVANDITASFTTSKEDMKEPNTT	378	
Dd	282	YLULSWK--LENRAYCPLOHQWTFDSITATECKSIHPVDEGTSSFVINTTVIGIELPDA	339	
Qy	379	YHLCNEBEEIAEKIKAVKNSTHSKYGD-LKYFKTDGGLYLVWQLIQNRLLDAKN--KL	435	
Dd	340	EKCICEOVNKHMEKVEAQDRVTKQEAITFITSGGALLAWLPLTPRLSATVKNLTREL	399	
Qy	436	NNETY-----RRSRRQAESTIDPMWGTMENGAGEYSSNSI	473	
Dd	400	TTPTSPPSPSPAPPASAARGSTPAALVLRERRRDAGNATP---VPPTAPGSKLGTLNP	456	
Qy	474	TVAQVOAYVDNLIRINNILEDSLKAECQHRAALVNWELSKIPTSMISINYRPVSA	533	
Dd	457	ATVQIFADVLSLRQRNMRLGDJARWCUEQQXRNQNVJRELKINPTTWMSGIYGAAVAA	516	
Qy	534	KRIDVISVINCIVDQTSVLSHKSLILLSADEKCFRSPPVTFPKFMDNSTIYKGOLGW	593	
Dd	517	KRLGDVISVSCQVPVNAQTVALRKSNR-VPGSETWCYSEPFLVSFINDTKTYEQOLD	575	
Qy	594	NEILLTTYLETQENTYYFOAKTDMYIYKHYEHKLVPLASSITTLDFTIALNFLLLEN	653	

```
QY 140 TTSTVYNGISDSTINQHTFYKSEIPWETEKMDTIYOCFNSRLNTGNNLTIVVDRDDIN 199
Db 104 TNILYNGWADSVTNREHEKFSVDSYETQMDTIYOCYNVAVKTDGLTRVYVDRDGVN 163
QY 200 MTVFLQPDVGTDPDKRYGQPELYLEPGFWGSYRRRTTVNCELMDMFARSNPPDFV 259
Db 164 ITVNLKPTGGLANGVRVYASQETELYDAPGMLIWTYRTTIVNCLITDMMAKSNPPDFV 223
QY 260 TATGDTVEMSPFSGEDDHENKMKHEKFWFVSINNYKVVDYQNRGTVPLGKTRIFLDREE 319
Db 224 TTTGQTVEMSPFVYDGNK--NKETFERADSFHVRTNYKIVDYNRGTNPOGERAFDLKGT 281
QY 320 YTLSEWHLKN--MSYCPCLTLWKAFYNGIOTHESSYHFVANDITASFTTSKEDMKENNT 378
Db 282 YTLSEWHLKN--LENRATYCPQLQHWOTFDSTIATETGKSIHFVTDGTSFVNTTVGIELPDA 339
QY 379 YHCLNEEIKAEIKKYAKVNSTHSHKYGD--LKYFKTDGGLYLVNQPLIQNRLLDKN--KL 435
Db 340 FKCIIEQVNTKHEKYEAVQDRYTKGQEAITYFITSGGLLLAWLPLTPRSLATVKNLTTEL 399
QY 436 NNETYSS-----RRSRQAESTTDPMMEMTNGAGGEYSSENSI 473
Db 400 TPTTSSPPSSPPAPSAARGSTPAAVLRRRRRDAGNATTP---VPPTAFKSLGLTLNPP 456
QY 474 TVAQVYADNLRIRINNILEDLSKAWCREQHRALVWNLKINPTSVMSIYNRPVSA 533
Db 457 ATVQIQFAYDSLRLQINRMGLDLARAWCLEQKQNMVLRRLTKINPTTVNMSIYKAVAA 516
QY 534 KRIGDVISVNCIIVVDQTSVSLHKSRLLSASDEKCFSPRPVTFKFMNDSTIYKQGLGVN 593
Db 517 KRLGDVISVQCVPNVQATVTLRKSMR--VPGSETMVCYSRPLVSFNFINDTKYEGQLGTD 575
QY 594 NEILLTTLTLETCQENTYFYFOAKTDMYIKYKVEHLKTVPLSSITLDTFIALNFTILEN 653
Db 576 NEIFLTKMTEVCQATSYFYQSGNEIHVYNDVHHFKTIELDGIATLQTFISLNTSLIEN 635
QY 654 VDFKVIETRDEKRLSNVFDIETPREYNYAQRVSLGRKLLD--LSTNRNOFVDAFGS 712
Db 636 IDFASLEYSRDEQRASNVFDLEGIFREYNFQAQNTIAGLRKLDNAVSGNRNOFVDGLGE 695
QY 713 LMDPLGAVGQTVVNAVSGVATLFSSIVTGTFINFKNPFQGMMLIIVVIGVLFAYFLTKK 772
Db 696 LMDSLGSGVGSITNLVSTVGLFSSVSGFISFPKPFQGMMLIIVLVAGVILVISLTR 755
QY 773 TKIYETAPIKWIPEIDKLKEREKSE---IAPISEELERIVLAMIHOONSHMETKR 829
Db 756 TRQMSQOPVQMLYPCIGIDELAQHASGEGPGINPISKTELOAIMLA--LHEQNOEQKRAAQ 813
QY 830 KDPKDSILTQAOQNLKR 847
Db 814 RAAGPSVASRALQAARD 831
```

RESULT 13

US-08-720-229-18

Sequence 18, Application US/08720229

Patent No. 6022542

GENERAL INFORMATION:

APPLICANT: Rose, Timothy M.

APPLICANT: Bosch, Marix L.

APPLICANT: Strand, Kurt

TITLE OF INVENTION: GLYCOPROTEIN B OF THE RHV/KSHV

SUBFAMILY OF HERPES VIRUSES

NUMBER OF SEQUENCES: 100

CORRESPONDENCE ADDRESS:

STREET: 755 Page Mill Road

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304-1018

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

```
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/720,229
FILING DATE: 26-SEP-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Schiff, J. Michael
REGISTRATION NUMBER: 40,253
REFERENCE/DOCKET NUMBER: 29938-20002.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 813-5600
TELEFAX: (415) 494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 857 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-720-229-18
```

Query Match 40.7%; Score 1848.5; DB 3; Length 857;

Best Local Similarity 46.1%; Pred. No. 2.9e-135; Mismatches 241; Indels 41; Gaps 12;

Matches 368; Conservative 148; Mismatches 241; Indels 41; Gaps 12;

QY 81 STPPYVVCASGVGVFRFQTDHVCPCD-ASDMVHSEGIILYKQNIIPFMFRVRYKRVV 139

Db 44 TSFPFVVCVLSHGDLFRSSDIQCPSCFCTRENHTGGLMVFKDNIIPYSFKVRSYTKIV 103

QY 140 TTSTVYNGISDSTINQHTFYKSEIPWETEKMDTIYOCFNSRLNTGNNLTIVVDRDDIN 199

Db 104 TNILYNGWADSVTNREHEKFSVDSYETQMDTIYOCYNVAVKTDGLTRVYVDRDGVN 163

QY 200 MTVFLQPDVGTDPDKRYGQPELYLEPGFWGSYRRRTTVNCELMDMFARSNPPDFV 259

Db 164 ITVNLKPTGGLANGVRVYASQETELYDAPGMLIWTYRTTIVNCLITDMMAKSNPPDFV 223

QY 260 TATGDTVEMSPFSGEDDHENKMKHEKFWFVSINNYKVVDYQNRGTVPLGKTRIFLDREE 319

Db 224 TTTGQTVEMSPFVYDGNK--NKETFERADSFHVRTNYKIVDYNRGTNPOGERAFDLKGT 281

QY 320 YTLSEWHLKN--MSYCPCLTLWKAFYNGIOTHESSYHFVANDITASFTTSKEDMKENNT 378

Db 282 YTLSEWHLKN--LENRATYCPQLQHWOTFDSTIATETGKSIHFVTDGTSFVNTTVGIELPDA 339

QY 379 YHCLNEEIKAEIKKYAKVNSTHSHKYGD--LKYFKTDGGLYLVNQPLIQNRLLDKN--KL 435

Db 340 FKCIIEQVNTKHEKYEAVQDRYTKGQEAITYFITSGGLLLAWLPLTPRSLATVKNLTTEL 399

QY 436 NNETYSS-----RRSRQAESTTDPMMEMTNGAGGEYSSENSI 473

Db 400 TPTTSSPPSSPPAPSAARGSTPAAVLRRRRRDAGNATTP---VPPTAFKSLGLTLNPP 456

QY 474 TVAQVYADNLRIRINNILEDLSKAWCREQHRALVWNLKINPTSVMSIYNRPVSA 533

Db 457 ATVQIQFAYDSLRLQINRMGLDLARAWCLEQKQNMVLRRLTKINPTTVNMSIYKAVAA 516

QY 534 KRIGDVISVNCIIVVDQTSVSLHKSRLLSASDEKCFSPRPVTFKFMNDSTIYKQGLGVN 593

Db 517 KRLGDVISVQCVPNVQATVTLRKSMR--VPGSETMVCYSRPLVSFNFINDTKYEGQLGTD 575

QY 594 NEILLTTLTLETCQENTYFYFOAKTDMYIKYKVEHLKTVPLSSITLDTFIALNFTILEN 653

Db 576 NEIFLTKMTEVCQATSYFYQSGNEIHVYNDVHHFKTIELDGIATLQTFISLNTSLIEN 635

QY 654 VDFKVIETRDEKRLSNVFDIETPREYNYAQRVSLGRKLLD--LSTNRNOFVDAFGS 712

Db 636 IDFASLEYSRDEQRASNVFDLEGIFREYNFQAQNTIAGLRKLDNAVSGNRNOFVDGLGE 695

QY 713 LMDPLGAVGQTVVNAVSGVATLFSSIVTGTFINFKNPFQGMMLIIVVIGVLFAYFLTKK 772

```
Db 696 LMDSLGVSQSTINLVSTVGGFSSLVSGFISFFKPNFGMLILVLVAGWILVISLRR 755
Qy 773 TKIYETAPIKMIYPIEDIKLKEGKSE---IAPISEEELERIVLAMHIHQONSHMETKTR 829
Db 756 TRMSQQPQVMLYPGIDELAQHASGEGPGINPISKTELQAIMLA--LHEQNOBKRAAQ 813
Qy 830 KDPKDSILTRAQNMRLKR 847
Db 814 RAAGPSVASRALQAARDR 831

RESULT 14
US-08-486-099-103
; Sequence 103, Application US/08486099
; Patent No. 6013263
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
; TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HEPATITIS
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,099
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-031
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 103:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 856 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-486-099-103

Query Match 40.4%; Score 1839; DB 3; Length 856;
Best Local Similarity 46.1%; Pred. No. 1.6e-134;
Matches 368; Conservative 148; Mismatches 240; Indels 42; Gaps 13;

Qy 81 STFPYRVCSAGVGDFVRFQTDHVCDD-ASDMVHSEGLIIYKQNIIPFMRVKRYKVV 139
Db 44 TSFPYRVCSLGHDLFRSSDIQCPSFGTRENHTHGLMWFKDNIIPYFVKVRSYTKIV 103
Qy 140 TTSYVNGIYSINQHTFYKSIPEWETEKMDTIYQCFNSLRNTGGNLLTYVDRDDIN 199
Db 104 TNLIIYNGYADSVNTRHEEKFSVDSYETDQMDTIYQCYNAVMTKDLTRYVDYRDGVN 163
```

```
Qy 200 MTFVLQPDVGVTDPVKRYGSOPELYLEPGFWGYSRRRTTVCNCLMDMFARSPFPDFV 259
Db 164 ITVNLKPTGGLANGVRRYASQTELYDAPGMLIWTYRTTVCNCLITDMMAKSNPFPDFV 223
Qy 260 TATGDTVEMSPFWGEGDDHKNKMEKFWFVSVINNYKVVDYQNRGTVPGLKTRIFLDREE 319
Db 224 TTTGQTVEMSPFYDGK--NKETPHERADSPHVRTNYKIVDYDNRGNTNPOGERRAFLOKGT 281
Qy 320 YTLSEWKLKN-MSYCPCLTLWKAFYNGIQTEHSGSYHFVANDITASFSTTSKEDMKEFNTT 378
Db 282 YTLNWK--LENRTAYCPLOHWQTFDSTIATETGSIHFVDEGTSSPVTNTVGIELPDA 339
Qy 379 YHCLNEEKAEIEKKYAKVNSTHSHKYGD-LKYFKTDGGLYLVWQPLIQNRLLDARN--KL 435
Db 340 FKCIIEQVN-KTHEKYEAVQDQRYTKGOEAIYFITSGLLLAWLPLTPRSLATVKNLTTEL 398
Qy 436 NNETY5-----RSRRQAESTTDDPMMENTGNGAGEVSSNSI 473
Db 399 TPTTSSPPSPSPAPSAARGSTPAAVLRRRRRRDAGNATTP--VPETAPGKSLGTLNLP 455
Qy 474 TVAQOVAYDNLRIINNILEDLSKWCROHRAALVMNELSKINPTSVMSIYNRPVSA 533
Db 456 ATVOIQFAYDSLRRQINRMGLDLARAWCLEOKRMVRLTKINPTITVMSIYKAVAA 515
Qy 534 KRIGDIVSVSNCIVVDQTSVSLHSLRLLSASDEKCFSPVPVTFKFMNDSTIYKQGLGVN 593
Db 516 KRLGDVLSVQCVPVQATVTLRKSMT-VPGSETWCYSRPLVSFSPFINDTKTYEGQLGTD 574
Qy 594 NEILLTTTYLETQCENTEYFQAKTDMYIYKNYBHLKTVPLSSITTLDTFIALNFTLLEN 653
Db 575 NEIFLTKMTVEVCATSQYVFGSNEIHVVNDYHFKTIELDGIATLQTFISLNTSLTEN 634
Qy 654 VDEKVIETYDEKRLSNVFDIETMREYNYVAQVSVCLRKDLDD-LSTNPNQVDFAGS 712
Db 635 IDFASLELYSDEQRASNVFDLEGIFREYNFQAQNIAGLRKDLONAVSNGRNQVDFGGE 694
Qy 713 LMDDLGAVGQTVNNAVSGVATLFSIVTGFNFINKNPPFGGMLMIIVIGVLFAIFYLTKK 772
Db 695 LMDSLGVSQGSITNLVSTVGGFSSLVSGFISFFKPNFGMLILVLVAGWILVISLRR 754
Qy 773 TKIYETAPIKMIYPIEDIKLKEGKSE---IAPISEEELERIVLAMHIHQONSHMETKTR 829
Db 755 TRMSQQPQVMLYPGIDELAQHASGEGPGINPISKTELQAIMLA--LHEQNOBKRAAQ 812
Qy 830 KDPKDSILTRAQNMRLKR 847
Db 814 RAAGPSVASRALQAARDR 830
```

RESULT 15

```
US-08-484-223B-103
; Sequence 103, Application US/08484223B
; Patent No. 6020459
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
; TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
; NUMBER OF SEQUENCES: 245
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
```


Qy	140	TTSTVYNGIYSDSI	NTQHTFYFKSI	BPWTEKMDTII	YQCFNSURLNT	TGGNLLTYVDRD	DDIN	199								
Db	104	TNIIILYNGWYADSV	TNRHEEKFSV	SDSETYDMDTII	YQCVNAVKMT	KDGLTRVYVDR	DGVN	163								
Qy	200	MTVFLOPVDGVT	PDVKRYGSOPEL	YLEGCFWGSYRR	TITUNCELMDM	FARSNPFPDFV	259									
Db	164	ITVNLKPTGG	LANGVRRYASQ	TELYDAPGLWLT	YTRTITVNC	LITDDMAKNSP	DFDV	223								
Qy	260	TATGDTVEMSP	FWSGEDDHENK	MKPWFVSINN	YKVVDYQNR	GTVPGLKTRIF	LDREE	319								
Db	224	TTTGOTVEMSP	FYDGK--NKET	FHERADS	FHVRTNYK	IVDYDNR	GTNQGERRAFLDKGT	281								
Qy	320	YTLSEWKHLK	N--MSYCP	PLTLKAFVNG	IQTEHSGSYH	FVANDITAS	FTTSTSKEDMK	EFNPT	378							
Db	282	YTLGK--	--LEN	RAYCP	LQHHQMTF	DSITATETG	KSIHFVTDG	TSFSVTNTV	GIELPDA	339						
Qy	379	YHCLNEEIKAE	IEKKYAKVN	STHSGYD--	LKYFKTD	GGLYLVN	QPLIQNRLLD	AKN--KL	435							
Db	340	FKCIEEQN--	KTHEK	YEAQDRTY	GQBAITFY	ITSGGLL	LAWLPLT	PRSLATV	KNLTEL	398						
Qy	436	NNETYS-----	-----RR	SRROA	ESTTDP	MWMTG	NGAGGEYS	SENSI	473							
Db	399	TTPTSSPPSP	PPAPSAAG	STPAVL	RRRRRD	AGNATP--	--VP	PTAPG	SLGT	LNNP	455					
Qy	474	TVAQVOYAYN	LRI	TRIN	NILED	SKACRE	QHRAALV	NELSKIN	PTS	VMSMIYNRP	VS	533				
Db	456	ATVQIQAYDS	LURROIN	RMLGDL	APAWC	LEQKRN	WLRLETK	INPTTW	MSIY	GKAVAA	515					
Qy	534	KRIGDVISN	CIVD	OTS	VSILH	LSRLLS	ASDEK	CFSPRP	PTPF	EMD	STIYKQ	LG	593			
Db	516	KRLGDVIS	VCVP	VNQAT	VL	LRKSMR--	VPGSET	WCYSR	PLVS	FIND	TKTYE	QGLGD	574			
Qy	594	NEILLTTYLET	COBNT	EYYFOAK	TDW	VIYKN	YHLKTV	PLSSIIT	TLDT	FIALN	FTLLEN	653				
Db	575	NEIFLTKWTE	VCAQ	TSQY	YFQSG	NEIHVYND	YHFKI	ELD	GIAT	LQTFIS	LNTSL	EN	634			
Qy	654	VDFKVI	ELYTRDE	KRLSN	VFDIET	MFREYNY	YAQRV	SLRK	DLDD--	LSTNR	NQFV	DAFGS	712			
Db	635	IDFAS	LELYSR	DEGRAS	NVFD	LEGIFREY	NFOAQ	NIAG	LK	DLN	AVSN	GRNQF	VDGLGE	694		
Qy	713	LMDDL	GAVGOTV	VNAV	SGVATL	FSSIV	TGFNFI	KNP	EGG	MLMI	IVIG	VLFAI	YFLPK	772		
Db	695	LMDS	SGSVQ	BSITNL	VTGGLF	SSLV	SGFIS	FFKN	P	EGG	MLIL	VLAG	WITVL	ISLTR	754	
Qy	773	TKIYET	APIK	MIYPEID	KLK	REGK	SE--	--IAP	ISEE	LERIV	LAM	HIHQ	NSHMET	KTR	829	
Db	755	TQMS	QQPVQ	MLYFCI	DELA	QHAS	GEGP	GINP	ISK	TEL	QAIM	LA--	LHEQ	NQ	QEKRAA	812
Qy	830	KDPK	DSIL	TR	QON	ML	RKR	847								
Db	813	RAAG	FSV	AS	R	AL	O	A	ADR	830						

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

Query Match 40.4%; Score 1839; DB 3; Length 856;

Qy	81	STTFPRVCSAGVGDFRFQTDHVCDD-ASDMVHSEGILLIYKONLIIPMFPRVRYKRVKV	139
Db	44	TSFPRVCELSSHGDLFRSSDIQCPSFGTRENHTEGLLMVFKNDIIPYSFKVRSYTKIV	103
Qy	140	TTSTVYNGIYSDSITNQHTFFKSTPEWETEKMDTIYCFNSLRNLGGNLLTTVYDRLDDIN	199
Db	104	TNLIYNGWYADSVNTRHEEFSYDSYETQMDTIYQYNAVKMTKDGLTRVYVDRDGVN	163
Qy	200	MTVFLOPVDGVTDPVKRYGSGOPELYLBPFGWGSYRRRTTVNCELMDMFAEASNPDPFFV	259
Db	164	ITVNLKPTGGGLANGVRIRYASOTELYDAPGLWIWTYRTRTVNCUITDMAKSNPSDPFFV	223
Qy	260	TATGDTVEMSPFSGEDDHENKMEKPMFVSINNYKVVDYQNGRTVPLGKTRIFLDREE	319
Db	224	TTTGQTVEMSPFYDGK--NKETPHERADSPHVRINYKIVDYNGRTNPOGERAEFLDKGT	281
Qy	320	YTLSEWEKHLKN-MSYCPBLTWKAFYNGIQTEHSGSHYFVANDITASFTTSKEDMKENFT	378
Db	282	YTLSNK--LENRTAYCPLOHQWQTFDSTIATETGKSIHFVTDEGTSFVNTTVTGIELPDA	339
Qy	379	YHCLNEEIKABIEKKYAVNSTHSGYGD-LKYPKTDGGLYLVWQPLIQNRLLDKKN--KL	435
Db	340	FKCIEQVN-KTHEKYAVODRYTKGEATYFTTSGGLLAWLPUPRSLATYKRNLT	398

RESULT 17
US-08-475-668A-103
; Sequence 103, Application US/08475668A
; Patent No. 6060065
; GENERAL INFORMATION:
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING INFLUENZA VIRUS
; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 211
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:

Qy 713 LMDLGVAGVTVNAVSGVATLFFSSIVGTGFINFKNPFPGMLMIIIVVIGVLFALYFLTKK 772
Db 695 LMDSLGSGVGSITNLVSTVGLFSSLVSGRISFFKPFNGMLLVLVAGVVLVISLTRR 754
Qy 773 TKIYETAPIKMIYPEIDKLKEREKSE---IAPISEEELERIVLAMHIHOONSHMETKTR 829
Db 755 TROMSQOPVQMLYPGIDELAQAQHASGEGPGINPISTKTELOAIMLA--LHEQNOEQKRAAQ 812
Qy 830 KDPKDSILTRAQNMLRKR 847
Db 813 RAAGPSVASRALQAARDR 830

RESULT 18
US-08-485-551A-103
; Sequence 103, Application US/08485551A
; Patent No. 6068973
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING INFLUENZA VIRUS
; NUMBER OF SEQUENCES: 211
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,551A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-023
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 103:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 856 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-485-551A-103

Query Match 40.48; Score 1839; DB 3; Length 856;
Best Local Similarity 46.13; Pred. No. 1.66-134;
Matches 368; Conservative 148; Mismatches 240; Indels 42; Gaps 13;

Qy 81 STPFYRVCSAGVGDFVFTDHDVCPD-ASDMVHSEGILLIYKONIIPFMRVRYKRVKVV 139
Db 44 TSPFPRVCELSSHGDLFRFSSDLCQSPGFRGNHTEGLLMVFKDNIIPYSFKVRSYTKIV 103
Qy 140 TTSTVYNGIYSDSITNOHTYKSIPEWETKMTDIYOCFNSRLNTGCGNLLTYVDRDIN 199
Db 104 TNIIITYNGWYADSVNRRHEEKFSVDSYETDQMDTIYOCYNNAVMTKDGRLTRVYVDRDGN 163

Qy 200 MTFVLOPVGVTEDVKRYGSOPELYLEPCGMWGSYRRRTTVNCELMDMPARSNPDPDFV 259
Db 164 ITVNLKPTGGLANGVRYASQTELYDAPGLWIWYTRITVNCIIITDMMAKSNPDPDFV 223
Qy 260 TATGDTVEMSPWMSGDDHENKMKHPWFVSVINNYKVVDYQNRGTVPGLKTRIFLDREE 319
Db 224 TTTGQIVEMSPFFDGK--NKETPHERADSFHVRTNYKIVDYDNRGTINPOGERAFLDKGT 281
Qy 320 YLTSWEKHLKN-MSYCPLTLWKAFYNGIOTHEHSGSYHFVANDITASTFTSKEDMKPEFTT 378
Db 282 YTLNWK--LENRTAYCPLQHWOTFDSTIATETGKSIHFVTDEGTSFVNTTGVIELPDA 339
Qy 379 YHCLNEEIKAEIEKYAKYNSTHSHKYGD-LKYFKTDGGLYLVWQPLIONRLLDKN--KL 435
Db 340 FCIEEQVN-KTHEKYEAQDRYTKGQEAITYFITSGLLLAWLPTPRSLATVKNLTEL 398
Qy 436 NNETYS-----RRRRQAESTTDPMMEMTNGCAGGEYSSENSI 473
Db 399 TPTTSPSPSPPPAPSAARGSTPAAVLRRRRDAGNATP---VPPTAGKSLGLTNNP 455
Qy 474 TVAQOYAYDNLRIRINNILEDLSKAWCREQHRAALVWNELSKINPTSVMSMIYNRPVSA 533
Db 456 ATVQIQFAYDSLRRQINRMGLDLARAWCLEQKRNQNVMLRELTKINPTVMSSTYGRKAA 515
Qy 534 KRIGDVISYNCIVVDOTSVSLHKSRLLSASDEKCFSPRPVTFKEMNDSTIYKGGOLGVN 593
Db 516 KRIGDVISYQCVVNVQATVLRKSMR-VPGSETMCSRPLVSFSFINDTKYEGQLGTD 574
Qy 594 NEILLTTTYLETQCENTEYFQAKTDMYIKNYEHLKTVPLSSITTLDTFIALNFTLEN 653
Db 575 NEIFLTKMTEVCQATSOYFQSGNEIHVYNDYHFKTELDGIAFLQFISLNTSLIEN 634
Qy 654 VDFKVIETREDEKLSNVFDIETMPREYNYAQRVSGLRKDLDD--LSTNRNQFVDAFGS 712
Db 635 IDFASLELYSRDEQRASNVFDLEGIPREYNFOAQNIAGLRKOLDNAVSNCRNQFVDCGLGE 694
Qy 713 LMDLGVAGVTVNAVSGVATLFFSSIVGTGFINFKNPFPGMLMIIIVVIGVLFALYFLTKK 772
Db 695 LMDSLGSGVGSITNLVSTVGLFSSLVSGIFSKNPPFGMLLVLVAGVVLVISLTRR 754
Qy 773 TKIYETAPIKMIYPEIDKLKEREKSE---IAPISEEELERIVLAMHIHOONSHMETKTR 829
Db 755 TROMSQOPVQMLYPGIDELAQAQHASGEGPGINPISTKTELOAIMLA--LHEQNOEQKRAAQ 812
Qy 830 KDPKDSILTRAQNMLRKR 847
Db 813 RAAGPSVASRALQAARDR 830

RESULT 19
US-08-471-913A-103
; Sequence 103, Application US/08471913A
; Patent No. 6093794
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING EPSTEIN-BARR VIRUS
; NUMBER OF SEQUENCES: 214
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711

QY	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
QY	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
QY	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
QY	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
QY	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
QY	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71</																													

```
Db 104 TNLIIYNGWYADSVNTRHEEKFSVDSYETQMDMTIYQCVNAVWMTKDGLTRVYVDRDGVN 163
Qy 200 MTVFLQVGVDPDVKRYGSOPELYLEPCFWGYSRRRTTVNCELMDMFARSPDPDFV 259
Db 164 ITVNLKPTGLANGVRRYASQTELYDAPGLIWTYRTRTTWNCILITDMMAKSNSPDPDFV 223
Qy 260 TATGDTVEMGPFWSGDEDDHKNKHEKFWFSVINNYKVVDYQNRGTVPGLKTRIFLDREE 319
Db 224 TTTGQTVEMSPFFYDGK--NKETHERADSFHVRTNYKIVDYDNRGTNPQGERAFLDKGT 281
Qy 320 YTLSEWKLKN-MSYCPLTLWKAFYNGIOTEHSGSYHFVANDITASFTTSKEDMKEFNTT 378
Db 282 YTLSSK--LENRTAYCPLQHWQTFDSTIATETGKSIHFVDEGTSSTVNTTVGIELPDA 339
Qy 379 YHCLNEEIKAEIEKKYAKVNSTHSGYD-LKYPKTDGGLYLVWQPLIQNRLLDAKN--KL 435
Db 340 FKCIIEQVN-KTHEKYEAVQDRYTKQGEAITFYFTSGGLLLAWLPLTPRSLATVKNLTEL 398
Qy 436 NNETY-----RRSRQAESTTDPMMEMTGNAGGEYSSENSI 473
Db 399 TPTSSPSPSPSPAPSAARGSTPAVLRRRRRDAGNATTP---VPPTAPGKSLGTLLNP 455
Qy 474 TVAQVQAYDNLRIIRINILEDSKAWCREQHRALVWNLKSKINPTSVMSIYNRPVSA 533
Db 456 ATVQIQFAYDSLRRQINRMGLDLARAWCLEQKQNMVLRRLTKINPTVMSSYIGRAVAA 515
Qy 534 KRIGDVISVNCIVDQTSVLSLHKSRLSASDEKCFSPRPVTFKFWNDSTIYKQGLGVN 593
Db 516 KRLGDVISVQCVPVNOATVTLRKSMR-VPGSETMCSYRPLVSPFINDTKYEGQLGTD 574
Qy 594 NEILLTTTLYLETCENTTEYFQAKTDMYIKVNYEHLKTVPLSSITTLDTFIALNFTLEN 653
Db 575 NEIFLTKKMTEVCOATSQYFQSGNEIHVNDYHFKTIELDGIATLQTFISUNTSLIEN 634
Qy 654 VDFKVIELTRDEKRLSNVFDIETMFREYNYAQRVSGLRKDLDD-LSTNRNQFVDAFGS 712
Db 635 IDPASLELYSRDEORASNVFDELEGIFREYNFQAQNIAGLRKDLDNVSNRGNQFVDDGLGE 694
Qy 713 LMDLGLAVGTQVNVANSGVATLSSIVTGFINFPKPFGGMLMIIIVIGVLFAIYFLTKK 772
Db 695 LMSLGSVGSGSITNLVSTVGGLFSSLSVGFISFPKPFGGMLILVLVAGVVIILVISLRR 754
Qy 773 TKIVETAPIKVIPEIDKLERECKSE--IAPISBEELERIVLAMIHOONGSHMETKTR 829
Db 755 TROMSQOPVOMLYPCIDELAQOASGEGPINPISKTELQAIMLA--LHEQNOEQKRAAQ 812
Qy 830 KDPKDSILTRAQNNMLRKR 847
Db 813 RAAGPSVASRALQAADR 830

RESULT 21
US-08-474-349A-103
; Sequence 103, Application US/08474349A
; Patent No. 633395
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING HUMAN PARAINFLUENZA
; TITLE OF INVENTION: VIRUS TRANSMISSION
; NUMBER OF SEQUENCES: 517
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
```

```
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,349A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-024
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 103:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 856 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-474-349A-103
```

```
Query Match 40.4%; Score 1839; DB 4; Length 856;
Best Local Similarity 46.1%; Pred. No. 1.6e-134;
Matches 368; Conservative 148; Mismatches 240; Indels 42; Gaps 13;

Qy 81 STPPYVCSASGVDFRFQTDHVCPCD-ASDMVHSGSGLLIYKONIIPMFVRVRYKRVV 139
Db 44 TSFPFVCSLSHGDDUFRSSDIQCFSGFGRNHTGLLMVFKDNIIPSGFVRSYTKIV 103
Qy 140 TTSVTVNGIYSISITNQHTFYKSIPEWETEKMDTIYQCFNSRLNTGNNLLTYVDRDIN 199
Db 104 TNLIIYNGWYADSVNTRHEEKFSVDSYETQMDTIYQCVNAVWMTKDGLTRVYVDRDGVN 163
Qy 200 MTVFLQVGVDPDVKRYGSOPELYLEPCFWGYSRRRTTVNCELMDMFARSPDPDFV 259
Db 164 ITVNLKPTGLANGVRRYASQTELYDAPGLIWTYRTRTTWNCILITDMMAKSNSPDPDFV 223
Qy 260 TATGDTVEMGPFWSGDEDDHKNKHEKFWFSVINNYKVVDYQNRGTVPGLKTRIFLDREE 319
Db 224 TTTGQTVEMSPFFYDGK--NKETHERADSFHVRTNYKIVDYDNRGTNPQGERAFLDKGT 281
Qy 320 YTLSEWKLKN-MSYCPLTLWKAFYNGIOTEHSGSYHFVANDITASFTTSKEDMKEFNTT 378
Db 282 YTLSSK--LENRTAYCPLQHWQTFDSTIATETGKSIHFVDEGTSSTVNTTVGIELPDA 339
Qy 379 YHCLNEEIKAEIEKKYAKVNSTHSGYD-LKYPKTDGGLYLVWQPLIQNRLLDAKN--KL 435
Db 340 FKCIIEQVN-KTHEKYEAVQDRYTKQGEAITFYFTSGGLLLAWLPLTPRSLATVKNLTEL 398
Qy 436 NNETY-----RRSRQAESTTDPMMEMTGNAGGEYSSENSI 473
Db 399 TPTSSPSPSPSPAPSAARGSTPAVLRRRRRDAGNATTP---VPPTAPGKSLGTLLNP 455
Qy 474 TVAQVQAYDNLRIIRINILEDSKAWCREQHRALVWNLKSKINPTSVMSIYNRPVSA 533
Db 456 ATVQIQFAYDSLRRQINRMGLDLARAWCLEQKQNMVLRRLTKINPTVMSSYIGRAVAA 515
Qy 534 KRIGDVISVNCIVDQTSVLSLHKSRLSASDEKCFSPRPVTFKFWNDSTIYKQGLGVN 593
Db 516 KRLGDVISVQCVPVNOATVTLRKSMR-VPGSETMCSYRPLVSPFINDTKYEGQLGTD 574
Qy 594 NEILLTTTLYLETCENTTEYFQAKTDMYIKVNYEHLKTVPLSSITTLDTFIALNFTLEN 653
Db 575 NEIFLTKKMTEVCOATSQYFQSGNEIHVNDYHFKTIELDGIATLQTFISUNTSLIEN 634
Qy 654 VDFKVIELTRDEKRLSNVFDIETMFREYNYAQRVSGLRKDLDD-LSTNRNQFVDAFGS 712
```

Db 635 IDFALELYSRDQRASNVDFLEGIREFYNFOAQNTAGRLKOLDNAVSNRQFVDGLGE 694
QY 713 LMDLGAQGVTVNAVSGVATLPSISIVTGFNFIKPNPFGMLMIIVIGVLAIFVLTAKK 772
Db 695 LMDSLGVSQITNLVSTVGGVLFSSLVSGFISFKPNPFGMLLVLVAGVILVLSLTKR 754
QY 773 TKIYETAPIKMIYPEIDKLEREGKSE---IAPISEEELERIVLAMHIHQONSHMETKTR 829
Db 755 TROMSQOPVQMLYPCIDELAQAQHASGEGPGINPISKTELQAIMLA--LHEQNEQOKRAAQ 812
QY 830 KDPKDSILTRAQNMRLKR 847
Db 813 RAAGPSVASRALQAARDR 830

RESULT 22
US-08-470-896-103
; Sequence 103, Application US/08470896
; Patent No. 6479055
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING HIV
; NUMBER OF SEQUENCES: 273
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,896
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-020
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/4864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 103:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 856 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-470-896-103

Query Match 40.4%; Score 1839; DB 4; Length 856;
Best Local Similarity 46.1%; Pred. No. 1.6e-134;
Matches 368; Conservative 148; Mismatches 240; Indels 42; Gaps 13;

QY 81 STFPYVCSASGVDVRFQTHVCPD-ASDMVHSEGILLIYKQNIIPMPFRVRYKVV 139
Db 44 TSPFPRVCSLSHGDLFRSSDIQCFSGFTRENHTGLMVPKDNIIIPYFVKRSYTKIV 103
QY 140 TTSTVYNGIYSDISITNQHTFYKSIPEWTERKMDTITQCFNSURLNTGGNLLTYVDRDDIN 199

Db 104 TNILYNGWYADSVTNRHEEKFSVDSYETDQMDTIYQYNVAKMTKDCGLTRVVDRCVN 163
QY 200 MTVFLOPVDGVTDPVKRYGSOPELYLEPCWFGWSYRRRTTVCBLMDMFARSNPPDFPV 259
Db 164 ITVNLKPTGGGLANGVRRYASQTELYDAPGMLIWTYRTRTTVNCCLITDMMAKSNPPDFPV 223
QY 260 TATGTVEMSPFWSGEDDHENKMEKPFVSVINNYKVVDYQNGRTVPLGTRIFLDREE 319
Db 224 TTTGQTVEMSPFYDGK--NKETFFERADS FVRYNYKIVDYNRGTNPQGERAFLDKGT 281
QY 320 YTLSEKHLKN--MSYCLPLTKAFYNGIQTEHSGSYHFVANDITASFTTSKEDKKEFNNT 378
Db 282 YTLGK--LENRTAYCPLQHWQTFDSTIATETGKSIHEVTDEGTSS FVTNTTVGIELPDA 339
QY 379 YHCLNEEIKAEIKYAKVNSTHSHKYGD-LKYFTDGLGLYLVQPLQNRLLDAKN--KL 435
Db 340 FKCIIEQVN-KTHEKYEAVQDRYTKGQEAITYFTTSGGLLLAWLPLTPRSLATVKNLTFL 398
QY 436 NNETYS-----RRSRQAESTTDPMMEMTGNAGGEYSSENSI 473
Db 399 TPTSSPPSPSPAPPASARGSTPAAVLRRRRRRDAGNATTP---VPPTAPKSLGTLNPP 455
QY 474 TVAQVQYAYDNLRIINNILEDLSKAMCREQRAALVMNELSKINPTSVMSMIYNRPVSA 533
Db 456 ATVOIQPAYDSLRRQINRMGLDLARAWCLEQKQNMVLRLELTKNPTTVMSSIVGKAVAA 515
QY 534 KRIGDIVSVSNQIVVDQTSVLSHLKSLLSASDEKCFRPPVTPKFMNDSTIYKGOLGVN 593
Db 516 KRLGDVIVSVQCVPNQATVTLRKSMT--VPGSETMCSYRPLVFSFINDTITYEGQLGTD 574
QY 594 NEILLTTLTYLETCOENTEYFQAKTDMYIKNYEHLKTVPLSSITTLDTFALNFTLLEN 653
Db 575 NEIFLTWKTEVCQATSOYIFQSGNEIHVYNDYHHFTIELDGIATLQTFISLNTSLTEN 634
QY 654 VDFKVIELYTRDEKRLSNVFDIETMFREYNYYAQRVSGRLRLDLD--LSTNRNQFVDFRGS 712
Db 635 IDFALELYSRDEQASNVDFLEGIREFYNFOAQNTAGRLKOLDNAVSNRQFVDGLGE 694
QY 713 LMDLGAQGVTVNAVSGVATLPSISIVTGFNFIKPNPFGMLMIIVIGVLAIFVLTAKK 772
Db 695 LMDSLGVSQITNLVSTVGGVLFSSLVSGFISFKPNPFGMLLVLVAGVILVLSLTKR 754
QY 773 TKIYETAPIKMIYPEIDKLEREGKSE---IAPISEEELERIVLAMHIHQONSHMETKTR 829
Db 755 TROMSQOPVQMLYPCIDELAQAQHASGEGPGINPISKTELQAIMLA--LHEQNEQOKRAAQ 812
QY 830 KDPKDSILTRAQNMRLKR 847
Db 813 RAAGPSVASRALQAARDR 830

RESULT 23
US-08-485-546A-103
; Sequence 103, Application US/08485546A
; Patent No. 6518013
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING EPSTEIN-BARR VIRUS
; NUMBER OF SEQUENCES: 214
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York

Db	635	IFASLELYSRDEQASNVFDLEGIFREYNFQAQNIAGLRKOLDNAVSNGRNQFVDGLGE	694
Qy	713	LMDDLGAVGQTVVNAVSGVATLFFSSIVTGFINFKNPGGMLIIVIGVLFAIVFLTK	772
Db	695	LMDSLGSVQGSITNLVSTVGGFLSSVSGFISFFKNPFGGMLILVLVAGVILVSLTR	754
Qy	773	TKIYETAFIKMIYPEIDKLKEREKSE---IAPISEEELERIVILAMHIHQONSHMETKTR	829
Db	755	TROMSQQPVMQLYPGIDELAAQHASGEGPGINPISKTELQAIMLA--LHEQNQOKRAAQ	812
Qy	830	KDPKDSILTRAQNMLRKR	847
Db	813	RAAGPSVASRALQAARDR	830
<p>RESULT 24</p> <p>US-08-804-439A-17</p> <p>Sequence 17, Application US/08804439A</p> <p>Patent No. 6015565</p> <p>GENERAL INFORMATION:</p> <p>APPLICANT: Rose, Timothy M.</p> <p>APPLICANT: Bosch, Marnix L.</p> <p>APPLICANT: Strand, Kurt</p> <p>TITLE OF INVENTION: GLYCOPROTEIN B OF THE RFHV/KSHV</p> <p>TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES</p> <p>NUMBER OF SEQUENCES: 113</p> <p>CORRESPONDENCE ADDRESS:</p> <p>ADDRESSEE: Fish & Richardson P.C.</p> <p>STREET: 4225 Executive Square, Ste 1400</p> <p>CITY: La Jolla</p> <p>STATE: CA</p> <p>COUNTRY: USA</p> <p>ZIP: 92037</p> <p>COMPUTER READABLE FORM:</p> <p>MEDIUM TYPE: Floppy disk</p> <p>COMPUTER: IBM PC compatible</p> <p>OPERATING SYSTEM: PC-DOS/MS-DOS</p> <p>SOFTWARE: Patent In Release #1.0, Version #1.30</p> <p>CURRENT APPLICATION DATA:</p> <p>APPLICATION NUMBER: US/08/804,439A</p> <p>FILING DATE: February 21, 1997</p> <p>CLASSIFICATION: 424</p> <p>ATTORNEY/AGENT INFORMATION:</p> <p>NAME: Haile, Lisa A.</p> <p>REGISTRATION NUMBER: 38,347</p> <p>REFERENCE/DOCKET NUMBER: 09176/004001</p> <p>TELECOMMUNICATION INFORMATION:</p> <p>TELEPHONE: (619) 678-5070</p> <p>TELEFAX: (619) 678-5099</p> <p>TELEX:</p> <p>INFORMATION FOR SEQ ID NO: 17:</p> <p>SEQUENCE CHARACTERISTICS:</p> <p>LENGTH: 849 amino acids</p> <p>TYPE: amino acid</p> <p>STRANDEDNESS: single</p> <p>TOPOLOGY: linear</p> <p>MOLECULE TYPE: protein</p> <p>US-08-804-439A-17</p>			
<p>Query Match 39.4%; Score 1792; DB 3; Length 849;</p> <p>Best Local Similarity 42.4%; Pred. No. 7.2e-131;</p> <p>Matches 354; Conservative 158; Mismatches 271; Indels 52; Gaps 13;</p>			
Qy	34	SPNTATWSTSPSLTGHYTHDSHGERGNNENRDSSEQNKNIYGSSTPPYRVCSASGV	93
Db	43	TPAQDAPTETPPPLS-----TNTNRGFEY-----FRVCGVAAT	75
Qy	94	GDVFFQTDHVCPOASDMVHSEGIILYKONIIPFERVVRKYRKVTTSTVYNGIYSDSI	153
Db	76	GETFRFDLCKTCPSTQDKKHVEGILLVTKINIVYIFKIRYRKIIITQLTWRGLTSSV	135
Qy	154	TNQHTFYKSIPEWTEKMDTIYQCFSNLRLNTGNLLTYVDRDDINMTVFLQPDVGTPD	213

```
Db 136 TGKFEATQAHEWEGDFDSIQVNSATMVNNRVQVYVDRDGVNKTINIRPVDGLTGN 195
Qy 214 VKRYGSOPELYLEPGWFGWYRRRTTVNCELMDMFARSNPPDFPVATGDTVEMSPFWS 273
Db 196 IQRYFSQPTLYSEPGMPGFYRVRTTVNCEIIVDMVARSMDPNYIATAGDLSLELSPQT 255
Qy 274 GEDDHENKMKHEK-PWFVSVINNYKVVDYQNRGTVPGLKTRIFLDREBYTSLWEKHLKMS 332
Db 256 FNTSQCTAPKRAMRVREVKNYKFVDYNNRGTAPAGQSRTELETPSATYSWKATATQTA 315
Qy 333 YCPETLWKAFYNGIQTEHSGSHYFVANDITAGFTTSKEDMKFNTTYHCLNEIEKAEIEK 392
Db 316 TCDLVHMKTFPRAIQTAHEHSYHFVANEVTATFNTPLEVENFTSYSCVSDQINKTISE 375
Qy 393 KYAKVNSTHSGYGLKYFKTDGGLYLVWQPLQNLRLDKNKLNNE-----TYSRRSRQ 447
Db 376 YIQKLNNSYVASKGYQFKTDGNLYLIWQPLEHPEIEDIDEDSDPEPTAPPKSTRKRE 435
Qy 448 AESTTDPMMEMTNGAGGEYSSENSITVAQVQYAYDNLRIRINNILEDLKAWCREQHRA 507
Db 436 AADNGNSTSEVS-----KGSENPLITAIQIOPAYDKLTTSVNNVLEELSRACREQVRD 488
Qy 508 ALVWNELSKINTSVMSIYNRPVSAKRIGDVISVNCIIVDQTSVSHKSLRLLSASDE 567
Db 489 TLMWYELSKVNPTSVMSAIYKGPVAAARYGDAISVTDICIYVDQSSVNIHQSLR-LQHDKT 547
Qy 568 KCFSRPPVTFKFMNDSTIYKGGOLGVNNEILLTTTVEITCOENTYFFQAKTDMYIKNYE 627
Db 548 TCYSRPRVTFKFINSTDPGLTQOLGPRKEIILSNINIECTKDESEHYFVGEYIYKNYI 607
Qy 628 HLKTVPLSSITTLDTFIALLNFTLENVDPKVELYTRDEKRL-SNVFDETFRFYNYYA 686
Db 608 FEEKLNLSIATLDTFIALLNISFIENIDFKTVELSSYTERKCLASSVFDIESMFRYNYT 667
Qy 687 QRVSLRKOLL-D-LSTNRNQFVDAFGSLMDLGAQVQTVVNAVSGVATLFFSSIVTGFINF 745
Db 668 YSLAGIKKDLNTIDYNRDLVQDLSDMADLIGRSVNVVSSVTFSSIVTGFIKF 727
Qy 746 IKNPFGGMLMIIVVIGLVFAIYFLTKTIYETAPIKMIYPEIDKLRKREG-----KSEIA 801
Db 728 FNPJGGFILLIIGIIFLVVLAARRNSQFHDPIKMLYPSVENYAAQAPPPYSASPP 787
Qy 802 PISEBELERIVLAME-IHQONSHMETK-TRKDPKDSILTRAQNMRLR-KRSGYSNL 853
Db 788 AIDKEIKRILLGMQVHQEKEAQKLTNSGP--TLWQKATGFLNRKKGISQL 840
```

RESULT 25

US-08-720-229-17
Sequence 17, Application US/08720229
Patent No. 6022542

GENERAL INFORMATION:

APPLICANT: Rose, Timothy M.
APPLICANT: Bosch, Marnix L.
APPLICANT: Strand, Kurt
TITLE OF INVENTION: GLYCOPROTEIN B OF THE RHFV/KSHV
SUBFAMILY OF HERPES VIRUSES

NUMBER OF SEQUENCES: 100

CORRESPONDENCE ADDRESS:

ADDRESSEE: Morrison & Foerster
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: CA

COUNTRY: USA

ZIP: 94304-1018

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/720,229

FILING DATE: 26-SEP-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Schifff, J. Michael
REGISTRATION NUMBER: 40,253
REFERENCE/DOCKET NUMBER: 29938-20002.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 813-5600
TELEFAX: (415) 494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 849 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-720-229-17

Query Match 39.4%; Score 1792; DB 3; Length 849;
Best Local Similarity 42.4%; Pred. No. 7, 2e-131;
Matches 354; Conservative 158; Mismatches 271; Indels 52; Gaps 13;

```
Qy 34 SPPNTATWSTSPLTGTHGYTHDSSHGERGNENRDSEQNKNIYGPSSTFFPYRVCSAGV 93
Db 43 TPAQDAPTETPPPLS-----TWTNRGFY-----FRVCGVAAT 75
Qy 94 GDVFRFQDTHVCPDASDMVHSEGLLIYKQNIIFPMFRVRKYRKVVTTSTVYNGIYSDSI 153
Db 76 GETFRFDLDKTCPSQDQKKHVEGILLVYKINIVPIFKIRRYRKIIITQITWRGLTSSV 135
Qy 154 TNOHTFYKSIPEWTEKMDTLYQCFNSLRLATGNLLTYVDRDDINMTVFLOPVDGVTPD 213
Db 136 TCKEMATQAHEWEGDFDSIYQCYNSATMVNNRVQVYVDRDGVNKTINIRPVDGLTGN 195
Qy 214 VKRYGSOPELYLEPGWFGWYRRRTTVNCELMDMFARSNPPDFPVATGDTVEMSPFWS 273
Db 196 IQRYFSQPTLYSEPGMPGFYRVRTTVNCEIIVDMVARSMDPNYIATAGDLSLELSPQT 255
Qy 274 GEDDHENKMKHEK-PWFVSVINNYKVVDYQNRGTVPGLKTRIFLDREBYTSLWEKHLKMS 332
Db 256 FNTSQCTAPKRAMRVREVKNYKFVDYNNRGTAPAGQSRTELETPSATYSWKATATQTA 315
Qy 333 YCPETLWKAFYNGIQTEHSGSHYFVANDITAGFTTSKEDMKFNTTYHCLNEIEKAEIEK 392
Db 316 TCDLVHMKTFPRAIQTAHEHSYHFVANEVTATFNTPLEVENFTSYSCVSDQINKTISE 375
Qy 393 KYAKVNSTHSGYGLKYFKTDGGLYLVWQPLQNLRLDKNKLNNE-----TYSRRSRQ 447
Db 376 YIQKLNNSYVASKGYQFKTDGNLYLIWQPLEHPEIEDIDEDSDPEPTAPPKSTRKRE 435
Qy 448 AESTTDPMMEMTNGAGGEYSSENSITVAQVQYAYDNLRIRINNILEDLKAWCREQHRA 507
Db 436 AADNGNSTSEVS-----KGSENPLITAIQIOPAYDKLTTSVNNVLEELSRACREQVRD 488
Qy 508 ALVWNELSKINTSVMSIYNRPVSAKRIGDVISVNCIIVDQTSVSHKSLRLLSASDE 567
Db 489 TLMWYELSKVNPTSVMSAIYKGPVAAARYGDAISVTDICIYVDQSSVNIHQSLR-LQHDKT 547
Qy 568 KCFSRPPVTFKFMNDSTIYKGGOLGVNNEILLTTTVEITCOENTYFFQAKTDMYIKNYE 627
Db 548 TCYSRPRVTFKFINSTDPGLTQOLGPRKEIILSNINIECTKDESEHYFVGEYIYKNYI 607
Qy 628 HLKTVPLSSITTLDTFIALLNFTLENVDPKVELYTRDEKRL-SNVFDETFRFYNYYA 686
Db 608 FEEKLNLSIATLDTFIALLNISFIENIDFKTVELSSYTERKCLASSVFDIESMFRYNYT 667
Qy 687 QRVSLRKOLL-D-LSTNRNQFVDAFGSLMDLGAQVQTVVNAVSGVATLFFSSIVTGFINF 745
Db 668 YSLAGIKKDLNTIDYNRDLVQDLSDMADLIGRSVNVVSSVTFSSIVTGFIKF 727
Qy 746 IKNPFGGMLMIIVVIGLVFAIYFLTKTIYETAPIKMIYPEIDKLRKREG-----KSEIA 801
```

Db 728 FTNPLGGIFILLIIGIIFLVVLLNRRNSQPHDAPIKMLYPSVNYAARQAPPYSASPP 787
Qy 802 PISEEELERIVLAMB-IHQONSHMETK-TRKDPKDSILTRAQNMLR-KRSGYSNL 853
Db 788 AIDKEEIKRILLGHQVHOEBEAKQOLTNNGP--TLWQKATGFLNRNRKGYSQL 840
RESULT 26
US-08-804-439A-15
; Sequence 15, Application US/08804439A
; Patent No. 6015565
; GENERAL INFORMATION:
; APPLICANT: Rose, Timothy M.
; APPLICANT: Bosch, Marnix L.
; APPLICANT: Strand, Kurt
; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RFHV/KSHV
; TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Ste 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: February 21, 1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 09176/004001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 678-5070
; TELEFAX: (619) 678-5099
; TELEX:
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 874 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-804-439A-15
Query Match 39.3%, Score 1787; DB 3; Length 874;
Best Local Similarity 43.5%; Pred. No. 1.8e-130;
Matches 367; Conservative 159; Mismatches 248; Indels 70; Gaps 20;
Qy 34 SPNTATWSTESPLTGHYGHDSHGGRNENRDSQKNIIYGSSTFPYRVCSA-SG 92
Db 58 SPQNT---STSKPSTDNQGSTPT-----IPTVDDTASKNFY-----KYRVCSSSS 102
Qy 93 VGDVFRFQTDHVCPSADSMVHSEGLILLYKQNIIPFMFRVKYRKVVTTSTVNGIYSDS 152
Db 103 SGELFRFDLQTCPTDKKXHVIGILLVKKNIIPYIFKVKYRKVIATSVTVYRGSQAA 162
Qy 153 ITNQHTYKIEPHEMETKMTIYCFNSLRNTGNNLLTVDRDDINMTVLPQVDGVTP 212
Db 163 VTNRDDISRAIPYNEISMIDRTYHCFSAMATVINGILNTYIDRDSENKSVLPQFVAGLTE 222
Qy 213 DVKRYGSOPELYLEPGHFWGSRRTTVNCELMDMPARSNPPEDFFVTATGDTVMSPFW 272
Db 223 NINRYFSQPLIYAEFGHFGPIYRVTTVNCVVDVMDYARSVEPYTHFTALGDTIEISPPC 282
Qy 273 SGEDDHENKM-----HEKPFVSVINNYKVVDYQNRGTVPGLKTRIFLDREYTL 322

Db 283 -----HNNSOCTTGNSTSRDATKVV---IBENHQTVDYERRGH-PTKDKRIFLKDEEYTI 333
Qy 323 SWEKHLKNMSYCPILTLWKAIFYNGIOIOTHSYSYHFVANDITASPTTSKEDMKEP--NT-TY 379
Db 334 SWKAEDRERAICDFVIMKTFPRAIOIITHNESPFHVANEVTSFLTNSQEEETELRGNTAIL 393
Qy 380 HCLNEEIKABIEKKYAKVNSTSHSKYGLKFKYFKTDGGLYLVQPLIQNRLLDKAKN---KLN 436
Db 394 NCWNSTINLEETVTKFNKSHIRDGEVKYKTKNGGLFLIWOAKMPLNLSHNTYTIERN 453
Qy 437 NETYSRRSRQAESTTDPMMEMTNGAGGYSSENSITVAQVOYAYDNLRIINILEDL 496
Db 454 NKTGNKSRQKRSVDT-----KTFQKAG-----LSTAQVOYAYDHLATSMNHILEEL 500
Qy 497 SKAWCREOHPAALVWNLKINPTSVMSMIYNRPVSAKRIGDVISVSNCIWVDTOSYSLH 556
Db 501 TKTWCREQKDNLMWYELSKINPVSVMAAIYGPVAVKMGDAFMVSECINVDQASYNIIH 560
Qy 557 KSLRLLSASDEK-CFSRPPVTFKPMNDSTIYKQGLGVNNEILLTTTLYLETCOENTEYFQ 615
Db 561 KSMR---TDDPKVCYSRPLVTFKAVNSTATFRQGLGTRNEILLTNTHTVETCRPTADHYFF 617
Qy 616 AKTDMYIKNYEHLKTVPLSSITLDTFTIALNFTLLENVDKVIYELTRDEKLSNVFDI 675
Db 618 VKMNTHYFKDYKFKVKTMDTNISTLDTFLTLNLTFTDIDFKTVELYSETERKVASALDL 677
Qy 676 ETMFREYNYAQRVSGLRKDL---LDLSTNRQPFDAFGLMDLDDLGAVGTGVNAVSGVA 732
Db 678 ETMFREYNYTQKLSLREDLNTIDL--NRDLVKDLSEMMADLDGIGKVVVNTFSGIV 735
Qy 733 TLFSSIVTGFINFIKNPPFGGLMIIVVIGVLPFIYFLTKTKIYETAPIKMIYPIEDKDK 792
Db 736 TVFGSIVGGFVSFTNPICGVITILLIIVVVFVIVSRRTNNNEAPIKMIYPIEDKDK 795
Qy 793 EREGKSEIAPISEEELERIVLAMHIIHQONSHMETKTRKDPK---SILTRAQNMLRKSG 849
Db 796 EQE---NIQPLPGEIKRILLGHQVHOEBEAKQOLTNNGP--TLWQKATGFLNRNRKGYSQL 840
Qy 850 YSNL 853
Db 852 YTRL 855

RESULT 27
US-08-720-229-15
; Sequence 15, Application US/08720229
; Patent No. 6022542
; GENERAL INFORMATION:
; APPLICANT: Rose, Timothy M.
; APPLICANT: Bosch, Marnix L.
; APPLICANT: Strand, Kurt
; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RFHV/KSHV
; TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/720,229
; FILING DATE: 26-SEP-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Schiff, J. Michael
; REGISTRATION NUMBER: 40,253

REFERENCE/DOCKET NUMBER: 29938-20002.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 813-5600
TELEFAX: (415) 494-0792
TELEX: 706141

```

, RELEX: 706141
, INFORMATION FOR SEQ ID NO: 15:
, SEQUENCE CHARACTERISTICS:
, LENGTH: 874 amino acids
, TYPE: amino acid
, STRANDEDNESS: single
, TOPOLOGY: linear
, MOLECULE TYPE: protein
US-08-720-329-15

```

Query Match	39.3%	Score 1787;	DB 3;	Length 874;
Best Local Similarity	43.5%	Pred. No. 1.8e-130;		
Matches	367;	Conservative 159;	Mismatches 248;	Indels 70; Gaps 20;
QY	34	SPPNATWSTESPLTGHVGHDSHGGRGNENRDPSEQNKNIYGSPTPYRVCSA-SG	92	
DB	58	SPQNT-----STSKPSTDNQGSTPT-----IPTVTDTASKNFY-----KYRVCSASS	102	
QY	93	VGVDFRQTDHVCPCDASDMVHSEGIILYKQNIIPFMFRVKRYKRVKVTSTVYNGIYSDS	152	
DB	103	SGELFRDLDOQCPDXTDKKHVGEILLVLKKNIVPIYFKVKRYKIATSTVYRGWSQAA	162	
QY	153	ITNQHTFYKSIPEWETEMKMTDIYOCFNSRLNTGNNLLTVYDRDDINMTVLPQVDGVTP	212	
DB	163	VTNRDDISRAIPYNEISMIDRTYHCFSAMATVINGILNTVIDRSENKSVLPQVAGLTE	222	
QY	213	DVKRYGSQPELYLEPGHFWGSYRRRTTVNCELMDMFARSPPDPFFVTATGDIVEMSPFW	272	
DB	223	NINRYFSQPLIYAEPPGPGIYRVRTTVNCEVDMYARSVEPYTHFITALGDTIEISPPC	282	
QY	273	SGEDDHENKM-----HEKPFVSVINNYKVQYQNRGTVPCLKTRIFLIDRBYTL	322	
DB	283	-----HNNSOCTTGNSTSRATKW-----IENHOTVDYIERGH-PTXDKRIFLKDEEYTI	333	
QY	323	SWEXHLKNMYSYCLTLWKAFYNGIQTBHSGSYHYFVANDITASFTTSKEDMKEF--NT-TY	379	
DB	334	SKAEDRERAIICDFVIWKTTPRAIQTIHNSFHFVANEVTASFLTSNQEETELRGNTIEL	393	
QY	380	HCLNEETKABIEKKYAKVNSTHSGYKGLKFKFTDGGLYLVQPLIQNRLDAXN---KLN	436	
DB	394	NCMNSTINETLEETVKKFNKSHIRDGKVYKTYNGGLFLIWAQMKPLNLSEHTNYTIERN	453	
QY	437	NETYSRRSRQAESTTDDPMEMTGNAGGEGVSSENSTVAQOVAYAYDNLAIRINNILEDL	496	
DB	454	NKTGNKSRQKRSVDT-----KTFQAGK-----LSTAQOVAYAYDHLRTSMNHLLEL	500	
QY	497	SKAWCREQRAALVWNBLSKINPTSVMSMYNRPVSAKRICDVISVNCITVVDQTSVSLH	556	
DB	501	TKTWCREQKDNLMWYELSKINPSVWAAIYGRFVAVKMGDAFWSSECINVDQASVNIH	560	
QY	557	KSURLLSASEK-CFSRPPVTFKPMNDSTIYKGLGVNNEILLTTTLYLETQCENTEYFYQ	615	
DB	561	KSMR-----TDDPKVCSYRPLVTFKFSNSTATFRGOLGTGRNEILLTNTHVETCRPTADHYF	617	
QY	616	AKTDMYIYKNYHLKTVPLSSIITLDPFIALNFTLLENVDFPKVIELYTRDEKRLSNVFDI	675	
DB	618	VKNWTHYFKDYKFKVKTMDTNISLTDTFLNLNLTFFIDNIDFKTYELSYETERKMASALDL	677	
QY	676	ETMFRYNYAYQRVSGLRKDL---LDLSTNRNQFVDAFGSLMDDLGAVGOTVNVNAVSGVA	732	
DB	678	ETMFRYNYTYQKLASREDLNDTIDL--NRDRLVKDLSEMMADLGDIGKVVNVTFSGIV	735	
QY	733	TLFSSIVTGFINFKPFGGMLMIIVVIGLVFAIYFLTKTKTIYETAPIKMIYPEIDKLK	792	
DB	736	TVFGSIVGGFVSFFTNPFIGGVTIILLIIVVVFVIVSRRRTNNMNEAPIKMIYPNIDKAS	795	
QY	793	EREKSEIAPISBEELBRIVLAMIHOONSHMETKTRDKPKD---SILTRAQNMLRKRSQ	849	
DB	796	EQE-----NIQPLGEEIKRILGHMQLOQSHGSEEBASHKPGFLFOLLGDLQLLRRR-G	851	

Db 119 EDL--DEGIMVVKRNIVAHFTKVRVQKVLFRSSVAYIHTTLLGSNTTEYVAPPWMEI 176
Qy 169 EKMDTIYOCNSRLNTGGNLLTYVDRDDI--NMTVFLQPVGDVTPDKRYKQSQPELYLEP 227
Db 177 HHINSHSQCYSSSRVVIAGTVFVAYHRDSYENKTMQLMDDSYNTHSTRTVTVKQDQHSR 236
Qy 228 GWFSGYRRRTTVNCELMDFARSNPPDFVATGDTVEMSPWGSDEDDHENKHEKWP 287
Db 237 GSTW--LYRETCNLCNMTIITARSKYPHFFATSTGVDVDISPPYNGTNRNASYFGENAD 295
Qy 288 FVSVINNYKVVDYQNRGTVPGLKTRI--FLDREYTLSE--KHLKNMSYCPPLTLWKAFYN 344
Db 296 KFFIFPNYTVSDGRENPALETHRLVAFLEADSVISWDIQDEKNT--CQLTFWEASER 354
Qy 345 GIOTHSYGVHVFANDITASFTTSKEDMKFNTYHCLNBEIKAEIEKKY--AKVNSTHSHK 403
Db 355 TIRSEADSYHFSKATWTATFLSKQEVNMSDALSDCVRDEAINKLOQIFNTSYNTQYK 414
Qy 404 YGDLKYFKTDGGLYLVWQPLIQNRLDANKLNNETYSRRSRQAESTTDPMMEMTNGA 463
Db 415 YGNVSVFETTGGLVFWQGIKQSLVELERLANRSSLNLTNRKST-----DGNNA 467
Qy 464 G--GEYSSENSITVAQVQYAYDNIRINNILEDLSKAWCREQHRAALVNELSKINPTS 521
Db 468 THLSNMESVHNLVYAQLQFTYDTRLGVINRALAQIAEAWCVDQRTLEVPFKLSKINPSA 527
Qy 522 VMSMIYNRPSAKRIGDVISVNCIVVDOTSVSLHLSRLLSASDEKCFRPPVTFKPMN 581
Db 528 ILSAIYNKPIAARFMDGLVGLASCVTINQTSVKVLRDMN--VKESPGRCYSRPVVIENFAN 586
Qy 582 DSTIYKQGLGVNNEILLTTTLYLETQENTTEYVFOAKTDMYIKNYEHLKTVPLSSITLD 641
Db 587 SSVYQYQGLGEDNELLGNHRTCEQLPSLKIFAGNSAYEYVDYLFKRMIDLSSTVD 646
Qy 642 TFIALNTFLNVDKVIYTRDEKRLSNVFDIETWFRYNYAQRVSGLRKDLDLST 701
Db 647 SMIALDIPLENTDFRVELYSQKELRSSNVDFLEIMREFNSYKQRYKVEDKVD-- 703
Qy 702 NRNOFVDFAGSLMDLCAVGQTVNVNAGVATLFPSSIVTGFINKPFGMLMIIIVIG 761
Db 704 PLPPYKGLDGLSLGGAAGAVGAIGAVGASVVEGATFLKNPFGAFTIILVAIA 763
Qy 762 VLFAYELTKTKTIYETAPIMXIPEI-----DKLKE-----R 794
Db 764 VIIIIYLYTRQRLCQPLQNLFPYLVASDGTIVTSGNTKQTSLOAPPSEYSEVNSGR 823
Qy 795 EG-----KSETAPISEBELERIVLAM-----HIHQONS-----HMETKTR--KPKOS 835
Db 824 KGPGPSDDASTAPPYTNQAYQMLLALVRLDAEQRAQNGTDSLDTGQTDQKQKPN 883
Qy 836 ILTPAQNLKRSYSLNKAESVE 860
Db 884 LLDLRLH---RKNGYRHLKSDSEE 905

RESULT 29

US-08-804-439A-19
; Sequence 19, Application US/08804439A
; Patent No. 6015565
; GENERAL INFORMATION:
; APPLICANT: Rose, Timothy M.
; APPLICANT: Bosch, Matnx L.
; APPLICANT: Strand, Kurt
; TITLE OF INVENTION: GYCOPROTEIN B OF THE RRVH/KSHV
; STREET: 4225 Executive Square, Ste 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESS: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Ste 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA

ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,439A
; FILING DATE: February 21, 1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 09176/004001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 678-5070
; TELEFAX: (619) 678-5099
; TELEX:
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 907 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-804-439A-19
Query Match 24.8%; Score 1128; DB 3; Length 907;
Best Local Similarity 30.9%; Pred. No. 3,9e-79;
Matches 286; Conservative 184; Mismatches 353; Indels 102; Gaps 25;
Qy 14 LWLYQVALYSLSIATG--VTSPNNTATWSTESPLTGHVGHDS--HGERGNENRDS 69
Db 5 IWCL--VVCUNLCIVCLGAASSTSRGTSATHS----HHSSHTTAAHSRSSVSQRTV 58
Qy 70 EQ-----NKNYIG-----SPSTPPYRVCASGSGVDVFRFOTDHC----PDA 108
Db 59 SSQTSHGVNETIYNTTLKVGVDVGVNTTKYPRVCMSAQGTDLIRFERNIVCTSMKPIN 118
Qy 109 SDMVHSEGLLIYKONIIPFMRVRYKRVKVTSTVYNGIYSDISTNOHTFFYKSIPEWET 168
Db 119 EDL--DEGIMVVKRNIVAHFTKVRVQKVLFRSSVAYIHTTLLGSNTTEYVAPPWMEI 176
Qy 169 EKMDTIYOCNSRLNTGGNLLTYVDRDDI--NMTVFLQPVGDVTPDKRYKQSQPELYLEP 227
Db 177 HHINSHSQCYSSSRVVIAGTVFVAYHRDSYENKTMQLMDDSYNTHSTRTVTVKQDQHSR 236
Qy 228 GWFSGYRRRTTVNCELMDFARSNPPDFVATGDTVEMSPWGSDEDDHENKHEKWP 287
Db 237 GSTW--LYRETCNLCNMTIITARSKYPHFFATSTGVDVDISPPYNGTNRNASYFGENAD 295
Qy 288 FVSVINNYKVVDYQNRGTVPGLKTRI--FLDREYTLSE--KHLKNMSYCPPLTLWKAFYN 344
Db 296 KFFIFPNYTVSDGRENPALETHRLVAFLEADSVISWDIQDEKNT--CQLTFWEASER 354
Qy 345 GIOTHSYGVHVFANDITASFTTSKEDMKFNTYHCLNBEIKAEIEKKY--AKVNSTHSHK 403
Db 355 TIRSEADSYHFSKATWTATFLSKQEVNMSDALSDCVRDEAINKLOQIFNTSYNTQYK 414
Qy 404 YGDLKYFKTDGGLYLVWQPLIQNRLDANKLNNETYSRRSRQAESTTDPMMEMTNGA 463
Db 415 YGNVSVFETTGGLVFWQGIKQSLVELERLANRSSLNLTNRKST-----DGNNA 467
Qy 464 G--GEYSSENSITVAQVQYAYDNIRINNILEDLSKAWCREQHRAALVNELSKINPTS 521
Db 468 THLSNMESVHNLVYAQLQFTYDTRLGVINRALAQIAEAWCVDQRTLEVPFKLSKINPSA 527
Qy 522 VMSMIYNRPSAKRIGDVISVNCIVVDOTSVSLHLSRLLSASDEKCFRPPVTFKPMN 581
Db 528 ILSAIYNKPIAARFMDGLVGLASCVTINQTSVKVLRDMN--VKESPGRCYSRPVVIENFAN 586
Qy 582 DSTIYKQGLGVNNEILLTTTLYLETQENTTEYVFOAKTDMYIKNYEHLKTVPLSSITLD 641

; MOLECULE TYPE: protein
PCT-US94-04180-2

```
Query Match      24.8%; Score 1128; DB 5; Length 907;
Best Local Similarity 30.9%; Pred.No.3.9e-79;
Matches 286; Conservative 184; Mismatches 353; Indels 102; Gaps 25;

QY 14 LWLYQVALYSLSIAETG--VTPSPNTATWSTESPLTGHYTHDSS--HGERGNNENRDS 69
DB 1 I WCL--VVCVNLICVCLGAAVSSSTRGTSATHS---HSSHTTSAHRSRSGSVQRVT 58
QY 70 EQ-----NKNIYG-----SPSTFPYRVCSASGVGDVRFOTDHYVC-----PDA 108
DB 59 SSQTVSHGNETIYNTTLKYGDVGWVNTTKYPRVCSMAQGTDLIRFERNICTSMKPIN 118
QY 109 SDMVHSEGLLIYKONIIPFMRKYRKVVTSTVYNGIYSDSIHQHTFYKSTPEWET 168
DB 119 EDL--DEGLMVYKRNIVAHTEFKVYQKVLTPRRSYAVIHTYLLGNSNTEYVAPPWEI 176
QY 169 EKMDTIYQCFNSRLRLNTGNNLLTYVDRODI--NMTVFLQPVGDVTPDKVRYGSOPELYLEP 227
DB 177 HHNSHSQCYSSYSRVIAGTVFVAYHRDSYENKTLQMLPDDYDSNTHSTRYTVTKDQWHSR 236
QY 228 GHWFGSVRRRTTVNCELMDMFARSNPPDPFVATGDTVEMSPFWSGEDDHNKMKHPW 287
DB 237 GSTW-LYRETCLNLCMTITITARSKYYPHFATSTGDDVVDISPFYNGTNRNASYFGENAD 295
QY 288 FVSVINNVKVDYQNRGTVPCLCKTRI--FLDREBYTLSWE-KHLKNMSYCPDLTKAFYN 344
DB 296 KFFIFPNYTVISDFGRPSNALSETHRLVAFERADSVISWDIQDEKNVT-CQLTFWEASER 354
QY 345 QIOTEGHSGSYHFVANDITASFTTSKEDMKFEFNTVHCLNEEIKAEIKKY-AKVNSTHSH 403
DB 355 TIRSEADSYHPSAKMTATFLSKQEVNMSDALDCVDEAKLQIIPNTSYNQYTEK 414
QY 404 YDGLKYFTDGLYLWQPLIQNRLLDAKNLNNTYRRSRQAEASTDDPMEMTNGA 463
DB 415 YGNVSVFETGGLVFWQGIQKSLVELERLANRSSNLTHNRKRT-----DGNA 467
QY 464 G--GEYSSENSITVAQVAYDNLRIRINNILEDLSKACREOHPAAVLWNELSKINTS 521
DB 468 THLSNMESVHNLYVAQLOFTYDLRGYNRLAAQIAEAWCVDQDRTELEVFKELSKINPSA 527
QY 522 VMSMIYNRPVSAKRIGDVISVNCIWIWDQTSVSLHSLRLISASDEKCFSPVPVTFKPMN 581
DB 528 ILSAIYNKPIARFWDVLGLASCTVINGTSVKVLRDMN-VKESPCRCYSRPFVFNAN 586
QY 582 DSTIYKQGLGVNNEILLTITTYLETQENTYFYFOAKTDMYIYKNYHKLTVPLSSITLD 641
DB 587 SSYVQVQGLGEDNEILLGNHRTTEECQPLSKIFIAGNSAYEYVDYLFKRMIDLSSISTVD 646
QY 642 TPIALNFTLLENVDPKVIETLYTRDEKRLSNVFDIETMFEYNYAQRVSGLKLKLLDST 701
DB 647 SMIALDIPLENTDFRVLLEYLSQKELRSNVFDLEEIMREFNSYKQRYVEDKVVD---703
QY 702 NRNQVDFAGSLMDLGLAGVQTVNNAVSGVATLFSSIVTGFNFINKFPFGMLIIVIG 761
DB 704 PLPPYLKGLDMLGSLGAGKAVGAIGAVGAVASVEGVATFLKNPFGAFTIILVAIA 763
QY 762 VLFAIYFLTKTKIYETAPIKMIYPEI-----DKLKE-----R 794
DB 764 VVIIIVLYTRQRLCMQPLQNLFPYLSADGTTVTSNKTDSLQAPPSVEESVYNSGR 823
QY 795 EG-----KSEIAPISSEELERIVLAM-----HIHQONS-----HMETKTR-KDPKDS 835
DB 824 KGPFPSSDASTAAPPYTNEQAYOMLLALVRLDAEQRAQQNGTSLDQGTGTQDQKQPN 883
QY 836 ILTRAQNMLRKSGYSNLKNAESVE 860
DB 884 LLDRLRH---RKNGYRHLKDSDEE 905
```

RESULT 33

```
US-08-220-151-9
; Sequence 9, Application US/08220151
; Patent No. 5529780
; GENERAL INFORMATION:
; APPLICANT: Paoletti, Enzo
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; TITLE OF INVENTION: CANINE HERPESVIRUS GB, GC AND GD AND USES THEREFOR
; NUMBER OF SEQUENCES: 91
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/220,151
; APPLICATION NUMBER: US/08/220,151
; FILING DATE: 30-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Prommer, William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2540
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; TELEX: 425066 CURTMS
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 906 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
; US-08-220-151-9
```

```
Query Match      24.7%; Score 1121.5; DB 1; Length 906;
Best Local Similarity 30.4%; Pred.No.1.2e-78;
Matches 282; Conservative 186; Mismatches 350; Indels 111; Gaps 25;

QY 14 LWLYQVALYSLSIAETGVTSPNTATWSTESPLTGHYTHDSSHGERGNN--ENRD---- 68
DB 5 I WCL--VVCVNLICVCLGAAVSSSSTSHATSS-----THNGSHTSRSTTSAQTRSVYSQ 55
QY 69 -----SEBQNKNIYG-----SPSTFPYRVCSASGVGDVRFOTDHYVC----- 105
DB 56 HVTSEAVSHRANETIYNTTLKYGDVGWVNTTKYPRVCSMAQGTDLIRFERNICTSMK 115
QY 106 PDASDMVHSGILLIYKONIIPFMRKYRKVVTSTVYNGIYSDSIHQHTFYKSTPE 165
DB 116 PINEDL--DEGLMVYKRNIVAHTEFKVYQKVLTPRRSYAVIHTYLLGNSNTEYVAPP 173
QY 166 WETEKMDTIYQCFNSRLRLNTGNNLLTYVDRODI--NMTVFLQPVGDVTPDKVRYGSOPELY 224
DB 174 WEIHHINKFAQCYSSYSRVIAGTVFVAYHRDSYENKTLQMLPDDYDSNTHSTRYTVTKDQW 233
QY 225 LEPQWFGSVRRRTTVNCELMDMFARSNPPDPFVATGDTVEMSPFWSGEDDHNKMKHE 284
DB 234 HSRGSTW-LYRETCLNLCMTITITARSKYYPHFATSTGDDVVDISPFYNGTNRNASYFGE 292
QY 285 KPWFSVINNVKVDYQNRGTVP--LCKTRIFLDREBYTLSWE-KHLKNMSYCPDLTKA 341
DB 293 NADKFFIFPNYTVISDFGRPNAAPEHRLVAFERADSVISWDIQDEKNVT-CQLTFWEA 351
QY 342 FYNGIOTEGHSGSYHFVANDITASFTTSKEDMKFEFNTVHCLNEEIKAEIKKY-AKVNST 400
```

Db 352 SERTIRSEADSYHFSSAKMTATATLSKQEVNMSDALSALDCVRDEAINKLOQIFNTSYNOT 411
Qy 401 HSKYGLDKYFKTDGGLYLWQPLQNLRLDANKLNNETYS---RRSRQAESTTDPME 457
Db 412 YEKYGNVSFTSGGLVFWGQIKQKSLVELEERANRSSLNITHRTSRSTSDNNTTHLSS 471
Qy 458 MTGNAGGEYSSENSITVAQVQYADNLIRINNILEDLSKAWCREQHRALVWNLSKI 517
Db 472 M-----ESVHNLVYAQLQFTYDYLGRYINRALAQIAEAWCVDORRTLEVFELSKI 522
Qy 518 NPTVMSMIYNRPVSAKRIGDVISVNCIVVDQTSVSLHKSRLLSASDEKCFSRPPTVF 577
Db 523 NPSAILSAIYNKPIAARFMDGLGLASCVTINQTSVKVLRDMN-VKESPCRCYSRPVWIF 581
Qy 578 KFMNDSTIYKQGVNNEILLTTTLETQENTYFQAKTDMYIYKNEYHLKTVPLSSI 637
Db 582 NFANSSVYQQLGEDNEILLGNHRTBECQPSLKIFAGNSAYEYVDYLFKRMIDLSI 641
Qy 638 TTLDTFIALNFTLLENVDFKVIELYTRDEKRLSNVFDIETMFRYNYAQRVGLRKDLL 697
Db 642 STVDMSIALDIDPLENTDFRVELYSQKELRSSNVFLEELMREFNSYKQKRVYEDKV 701
Qy 698 DLSTNRNQVDAFGLMDLGAAGQTVVNAVSGVATLFSISVTGFINFKNPFGLMII 757
Db 702 D---PLPPYLKGLDLDLMSGLGAAGKAVGAIGAVGAVASVVEGVATFLKNPFGAFTIL 758
Qy 758 VVIGVLFALYFLTKTKIYETAPIKMIYPEI-----DKLKE----- 793
Db 759 VAIAVVIITYLITRQRLCTQPLQNLFPYLVLSADGTTVTSGSTKOTSLQAPPSYESVY 818
Qy 794 ---REG-----KSEIAPISEEELERIVLAM-----HIHQONS-----HMETKTR-KD 831
Db 819 NSGRKGCPSSDASTAAPPTNQAYQVQMLLALRLDAEQAQNGTDSLDTGQTQDKG 878
Qy 832 PKOSILTRAQNMRLKRGYSNLKNAESVE 860
Db 879 QKPNLLDLRLH---RKNGYRLKXDSDEE 904

RESULT 34
US-08-413-118-9
; Sequence 9, Application US/08413118
; Patent No. 568920
; GENERAL INFORMATION:
; APPLICANT: PAOLETTI, ENZO
; APPLICANT: LIMBACH, KEITH J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; TITLE OF INVENTION: CANINE HERPESVIRUS GB, GC, AND GD AND USES THEREFOR
; NUMBER OF SEQUENCES: 128
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
; STREET: 530 FIFTH AVENUE, 25TH FLOOR
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/413,118
; FILING DATE: 29-MAR-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/220,151
; FILING DATE: 30-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: FROMMER, WILLIAM S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2670
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 906 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
; US-08-413-118-9

Query Match 24.7%; Score 1121.5; DB 1; Length 906;
Best Local Similarity 30.4%; Pred No 1.2e-78;
Matches 282; Conservative 186; Mismatches 350; Indels 111; Gaps 25;

Qy 14 LMWLYQVALYSLSIABTGTVPSPNTATWSTESPLTGHYTHDSSHGGRNN-ENRD--- 68
Db 5 IWCL--VVCNLICIVCLGAAVSSSTSHATSS-----THNGSHTSRITSAQTRSVYSQ 55
Qy 69 -----SEEQNKNIY-----SPSTFPYRVCSASGVDFRQTDHVC----- 105
Db 56 HVTSEAVSHRANETIYNTTLKYGDVVGVTNTTKYRVCSMAQGTDLIRFERNIICTSMK 115
Qy 106 PDASDMVHSEGIILLIYKONIIPEMPRVKRVKVTSTVYNGIYSDSIITQHTFYKSIIEP 165
Db 116 PINEDJ--DEGMVYKRNIVAHFKRVYQKVLTPRRSYAYIYTYLLGSENTEYVAPPM 173
Qy 166 WETEKMDTIYQCNSURLNTGGNLLTYVDRDDI--NMTVFLQPVGDVTPDVKRYGSOPELY 224
Db 174 WEIHINKPAQCYSYRVIGGTVPVAYHRDSYENKTMQLIPDDYNTSTHSTRVTVTKDQW 233
Qy 225 LEPGFWGYSRRRTTVNCELMDFARSNPPFPFVATGDTVEMSPFWGDEDDHKNMHE 284
Db 234 HSRGSTW-LYRETCLNCLMTITARSKYPIHFATSTGDDVYVISPFGYNGTNRNASYFGE 292
Qy 285 KPMFVSVINNYKV--DYQNRGTVP-LGKTRIIFLDREYTLSE--KHLKMSYCPLTLWKA 341
Db 293 NADKFFIFNYTIIVSDFGRNAPETHRLVAFLEADSVISWDIQDEKNT-COLTFWEA 351
Qy 342 FYNGIQTEHSGSYHFVANDITASTFTSKEDMKEFTTYHCLNEEIKAEIEKKY-AKVNST 400
Db 352 SERTIRSEADSYHFSSAKMTATATLSKQEVNMSDALSALDCVRDEAINKLOQIFNTSYNOT 411
Qy 401 HSKYGLDKYFKTDGGLYLWQPLQNLRLDANKLNNETYS---RRSRQAESTTDPME 457
Db 412 YEKYGNVSFTSGGLVFWGQIKQKSLVELEERANRSSLNITHRTSRSTSDNNTTHLSS 471
Qy 458 MTGNAGGEYSSENSITVAQVQYADNLIRINNILEDLSKAWCREQHRALVWNLSKI 517
Db 472 M-----ESVHNLVYAQLQFTYDYLGRYINRALAQIAEAWCVDORRTLEVFELSKI 522
Qy 518 NPTVMSMIYNRPVSAKRIGDVISVNCIVVDQTSVSLHKSRLLSASDEKCFSRPPTVF 577
Db 523 NPSAILSAIYNKPIAARFMDGLGLASCVTINQTSVKVLRDMN-VKESPCRCYSRPVWIF 581
Qy 578 KFMNDSTIYKQGVNNEILLTTTLETQENTYFQAKTDMYIYKNEYHLKTVPLSSI 637
Db 582 NFANSSVYQQLGEDNEILLGNHRTBECQPSLKIFAGNSAYEYVDYLFKRMIDLSI 641
Qy 638 TTLDTFIALNFTLLENVDFKVIELYTRDEKRLSNVFDIETMFRYNYAQRVGLRKDLL 697
Db 642 STVDMSIALDIDPLENTDFRVELYSQKELRSSNVFLEELMREFNSYKQKRVYEDKV 701
Qy 698 DLSTNRNQVDAFGLMDLGAAGQTVVNAVSGVATLFSISVTGFINFKNPFGLMII 757
Db 702 D---PLPPYLKGLDLDLMSGLGAAGKAVGAIGAVGAVASVVEGVATFLKNPFGAFTIL 758
Qy 758 VVIGVLFALYFLTKTKIYETAPIKMIYPEI-----DKLKE----- 793
Db 759 VAIAVVIITYLITRQRLCTQPLQNLFPYLVLSADGTTVTSGSTKOTSLQAPPSYESVY 818
Qy 794 ---REG-----KSEIAPISEEELERIVLAM-----HIHQONS-----HMETKTR-KD 831

```
Db 819 NSGRKGPSSDASTAAPPYNEQAYQMLLALRLDAEQRAQQNGTDSLDCQTQDKG 878
Qy 832 PKDSILTRAQNMRLKRSYGNLKNKNAESVE 860
Db 879 QKPNLLDLRLH---RKNGYRHLKDSDEE 904

RESULT 35
US-08-473-446-9
; Sequence 9, Application US/08473446
; Patent No. 6017542
; GENERAL INFORMATION:
; APPLICANT: PAOLETTI, ENZO
; APPLICANT: LIMBACH, KEITH J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; TITLE OF INVENTION: CANINE HERPESVIRUS gb, gc, and gd AND USES THEREFOR
; NUMBER OF SEQUENCES: 128
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
; STREET: 530 FIFTH AVENUE, 25TH FLOOR
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/473,446
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/413,118
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: FROMMER, WILLIAM S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2670
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 906 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
US-08-473-446-9

Query Match 24.7%; Score 1121.5; DB 3; Length 906;
Best Local Similarity 30.4%; Pred. No. 1.2e-78;
Matches 282; Conservative 186; Mismatches 350; Indels 111; Gaps 25;

Qy 14 LWLYQVALYSLSIATGVTGPPNTATWSTSPLTGCHYTHDSSHGEGNN-ENRD---- 68
Db 5 IWCL--VVCVNLICVCLGAAVSSSSSTSHATSS-----THNGSHTSRTTSAQTRSVYSQ 55
Qy 69 -----SEQGNKIYG-----SPSTFPYRVCASGVGDVFRFQTDHVC----- 105
Db 56 HVTSEAVSHRANETIYNTTKYGVGVGNTTKYPRVCSMAQGTDLIRFERNICTSMK 115
Qy 106 PDASDMVHSEGLIYKNIIPFMRVRKRYKVVTTSTVYNGIYSDSITNQHTFYKSTEP 165
Db 116 PINEDL--DEGLMVVYKRNIVAHTEKRVYQKVLTFRRSYAYIYTYLLGNTETVAPPM 173
Qy 166 WETEKMDTYQCFNSLRNLNTGNLLTYVDKDI-NMTVFLOPVDGVTDPVKRYGSGPELY 224
Db 174 WEIHHINKFAQCYSYSRVIQGVTFVAVHRDSYENKTMQLIPDDYSNTHSTHYTVTKDQW 233
```

```
Qy 225 LEPQWFGSYRRRTTVCNCELMDFARNPPDFVFVATGDTVEMSPFWSGDDHDKMHE 284
Db 234 HSRGSTW-LYRETCLNACMLTITARSKYPHFATSTGDDVYISPFYNGTNRNASYGE 292
Qy 285 KPWFSVSVNNYKV-V-DYONRGTV-P-LGKTRIFLDREYTL-SWE-KHLKNMSYCPCLTLKA 341
Db 293 NADKFFIFPNYTVISDFGRPNAAPETHRLVAFLEADSVISWDIQDEKNVT-CQLTFWEA 351
Qy 342 FYNGIOTHEHSGSYHFVANDITASFTTSKEDMKENTTYHCLNEIEIKAEIEKKY-AKNST 400
Db 352 SERTIRSEADESYHFSAKMTATFLSKQEVNMSDSDALDCVRDEAINKLOQIFNTSYNQ 411
Qy 401 HSKYGDLYKFKTDGGLYLVQPLIQNRLDLDAKNLNNETYS---RRSRROAESTTDPNME 457
Db 412 YEKYGNVSFVETSGGLVVFQGIQKSLVELERLANRSSLNITHTRRSISDNNTTHLSS 471
Qy 458 MTGNAGGEVSSENSITVAQVQYAYDNLIRININILEDLSKAWCREQHRAALVNEISKI 517
Db 472 M-----ESVHNLVYAQLQFTYDTRLGYINRALAQIAEAMCVDQRRTLFVFKELSKI 522
Qy 518 NPTSVMSMIYNRPVSAKRIGDIVSVNCIVVDOTSVSLHKSLLSLSASDEKCFRPPVTF 577
Db 523 NPSAILSAIYNKPIAARFMGVDGLASCTVINTQTSVKVLRDMN-VKESPGRCYSRPVIF 581
Qy 578 KFMNDSTIYGQLGQVNNELITLTYYLETQCOENTEVYFOAKTDMYIKNYEHLKTVPLSSI 637
Db 582 NFANSSVVOYQGLGEDNEILLGNHRTBECQPSUKIFIAGNSAYEYVDYLFKRMIDLSSI 641
Qy 638 TTLDTFIALNFTLLENVDVKVIELYTRDEKRLSNVFDIETMFRYNYVAORVSGLRKDLL 697
Db 642 STVDSMIALDIDPLENTDFRLELYSQELRSSNVFDEEIMREBFSYKQRYKYVEDKV 701
Qy 698 DLSTNRNQFVDAFGSLMDDLGAVCQTVNAVSGVATLFSISIVTGFIKPKFPGMLMI 757
Db 702 D---PLPPYLKGLDGLMSGLGAAGKAVGAIGVAGVAVSVEGVATFLKNPFGAFTIL 758
Qy 758 VWIGVFAIYPLTKTKIYETAPIKMIYPEI-----DKLKE----- 793
Db 759 VAIAVIITYLIYTRQRLCTQPLNQNLFPYLVSDAGTTVTSGSTKDTSLQAPPSYEESVY 818
Qy 794 ---REG-----KSEIAPISEEELERIVLAM-----HIHQONS-----HMETKTR-KD 831
Db 819 NSGRKGPSSDASTAAPPYNEQAYQMLLALRLDAEQRAQQNGTDSLDCQTQDKG 878
Qy 832 PKDSILTRAQNMRLKRSYGNLKNKNAESVE 860
Db 879 QKPNLLDLRLH---RKNGYRHLKDSDEE 904
```

```
RESULT 36
US-08-804-439A-20
; Sequence 20, Application US/08804439A
; Patent No. 6015565
; GENERAL INFORMATION:
; APPLICANT: Rose, Timothy M.
; APPLICANT: Bosch, Marnix L.
; APPLICANT: Strand, Kurt
; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RPHV/KSHV
; TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Ste 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
```

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,439A
; FILING DATE: February 21, 1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 09176/004001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 678-5070
; TELEFAX: (619) 678-5099
; TELEX:
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 830 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-804-439A-20

Query Match      23.1%; Score 1050.5; DB 3; Length 830;
Best Local Similarity 32.3%; Pred. No. 3,6e-73;
Matches 244; Conservative 160; Mismatches 293; Indels 59; Gaps 15;

QY      83 FPRVCSAGVGDFRFQTDHVC-PDASDMVHSEGIILYKQNIIPMFVRVRYKRVKVT 141
Db      36 YPRICSIAGTDLMRDRDISCPYKSNAMSGFFIIYKTIETIYTFPVRYKKELTF 95
QY      142 STVNGIYSDSITNQHTFYKSIETPEWETEKMDTIYQCFNSRLN--TGGNLLTYVDRDIN 199
Db      96 QSSYRDVGVVYFLDRTVMGLAMPVYEANLVNSHAQCYSAVAMKRPDGTVFSAFHEDNNKN 155
QY      200 MTVPLOPVGVTDPVKRYGSOPELYLEPGFWGYSRRRTTVNCELMDMPARSNPPDFV 259
Db      156 NTNLNPLNFKSITNRKFTITKEPYFARGPLW-LYSTSTSLNCIVTEATKAKYPSYFA 214
QY      260 TATGDTVEMSPWGSDE-DHENKMHKPMFVSVINNYKV--RGTVPLGKTRIFL 315
Db      215 LTTGEIVGSGPPNGSNGKHFAEPLK--LTTILENTWIEDLMNGMNGATTLLVRKIAPL 271
QY      316 DREYTLSEKHLKNNMSYCPDLTWKAFYNGIQTEHSGSYHFVANDITASTTTSKEDMKEF 375
Db      272 EKADTLFSWEIKEENESVCMLEKHTTTVTHGLRAETDTHYFISKELTAAFVAPKESLNT 331

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/720,229-20
; Sequence 20, Application US/08/720229
; Patent No. 6022542
; GENERAL INFORMATION:
; APPLICANT: Rose, Timothy M.
; APPLICANT: Bosch, Marnix L.
; APPLICANT: Strand, Kurt
; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RPHV/KSHV
; TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/720,229
; FILING DATE: 26-SEP-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Schiff, J. Michael
; REGISTRATION NUMBER: 40,253
; REFERENCE/DOCKET NUMBER: 29938-20002.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 830 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-720-229-20

Query Match      23.1%; Score 1050.5; DB 3; Length 830;
Best Local Similarity 32.3%; Pred. No. 3,6e-73;
Matches 244; Conservative 160; Mismatches 293; Indels 59; Gaps 15;

QY      83 FPRVCSAGVGDFRFQTDHVC-PDASDMVHSEGIILYKQNIIPMFVRVRYKRVKVT 141
Db      36 YPRICSIAGTDLMRDRDISCPYKSNAMSGFFIIYKTIETIYTFPVRYKKELTF 95
QY      142 STVNGIYSDSITNQHTFYKSIETPEWETEKMDTIYQCFNSRLN--TGGNLLTYVDRDIN 199
Db      96 QSSYRDVGVVYFLDRTVMGLAMPVYEANLVNSHAQCYSAVAMKRPDGTVFSAFHEDNNKN 155
QY      200 MTVPLOPVGVTDPVKRYGSOPELYLEPGFWGYSRRRTTVNCELMDMPARSNPPDFV 259
Db      156 NTNLNPLNFKSITNRKFTITKEPYFARGPLW-LYSTSTSLNCIVTEATKAKYPSYFA 214
QY      260 TATGDTVEMSPWGSDE-DHENKMHKPMFVSVINNYKV--RGTVPLGKTRIFL 315
Db      215 LTTGEIVGSGPPNGSNGKHFAEPLK--LTTILENTWIEDLMNGMNGATTLLVRKIAPL 271
QY      316 DREYTLSEKHLKNNMSYCPDLTWKAFYNGIQTEHSGSYHFVANDITASTTTSKEDMKEF 375
Db      272 EKADTLFSWEIKEENESVCMLEKHTTTVTHGLRAETDTHYFISKELTAAFVAPKESLNT 331

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,439A
; FILING DATE: February 21, 1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 09176/004001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 678-5070
; TELEFAX: (619) 678-5099
; TELEX:
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 830 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-804-439A-20

Query Match      23.1%; Score 1050.5; DB 3; Length 830;
Best Local Similarity 32.3%; Pred. No. 3,6e-73;
Matches 244; Conservative 160; Mismatches 293; Indels 59; Gaps 15;

QY      83 FPRVCSAGVGDFRFQTDHVC-PDASDMVHSEGIILYKQNIIPMFVRVRYKRVKVT 141
Db      36 YPRICSIAGTDLMRDRDISCPYKSNAMSGFFIIYKTIETIYTFPVRYKKELTF 95
QY      142 STVNGIYSDSITNQHTFYKSIETPEWETEKMDTIYQCFNSRLN--TGGNLLTYVDRDIN 199
Db      96 QSSYRDVGVVYFLDRTVMGLAMPVYEANLVNSHAQCYSAVAMKRPDGTVFSAFHEDNNKN 155
QY      200 MTVPLOPVGVTDPVKRYGSOPELYLEPGFWGYSRRRTTVNCELMDMPARSNPPDFV 259
Db      156 NTNLNPLNFKSITNRKFTITKEPYFARGPLW-LYSTSTSLNCIVTEATKAKYPSYFA 214
QY      260 TATGDTVEMSPWGSDE-DHENKMHKPMFVSVINNYKV--RGTVPLGKTRIFL 315
Db      215 LTTGEIVGSGPPNGSNGKHFAEPLK--LTTILENTWIEDLMNGMNGATTLLVRKIAPL 271
QY      316 DREYTLSEKHLKNNMSYCPDLTWKAFYNGIQTEHSGSYHFVANDITASTTTSKEDMKEF 375
Db      272 EKADTLFSWEIKEENESVCMLEKHTTTVTHGLRAETDTHYFISKELTAAFVAPKESLNT 331
```


Qy	376	NTTYHCLNEEBIKABIEKKY-AKVNSTHSKYGDLYKFKTDDGGIYLVWQPLIQNRLLDAXNK	4334
Db	332	DPKOTCIKDEPEKIIINEYMSDYNDYTSVMGNSYQIFKTTGDLILIQVLQVKSMLFL---	388
Qy	435	LNNETYSRRSRQRAESTTDPMMEMTNGCAGGEYSSENSITVAQVOYAYVDNLRIRINNILE	494
Db	389	---EQGSEKIRRRRDV-----DVKSRHDILYVQLQYLDYTLKDYINDALG	431
Qy	495	DL\$KAWCREQHRHAALVWNEL\$KINPT\$VMSMIYNRPV\$AKRIG\$DVISV\$NCIN\$VWDQTSV\$	554
Db	432	NLAESWCLDQKRTITMLHEL\$KISPS\$VSEYVGRPI\$SAQLHG\$DVLAI\$KSCIEV\$NQSSVQ	491
Qy	555	LHKSLRLLSA---SDEKCF\$RPPVTFKFMDS-TIYKGLQGVNNEIILLTTLYLETCCQEN	609
Db	492	LHKS\$RVVD\$AKVSR\$ETWCY\$NRP\$LVTF\$FV\$NSTPEV\$FCQGL\$DNEIILLG\$DHRT\$EBCIP	551
Qy	610	TEYTFQAKTDMYIYKNVEHLKTVPLSSITTLDTFIALNFTLLENVDFKVIELTYTRDEKRL	669
Db	552	STKIFL\$GNH\$AVVTDYTH\$TNSTPIEDIEV\$DAFIRL\$KIDPLENAD\$FVLDV\$SPDELSR	611
Qy	670	SNVFEDIETMREYNVY\$Q\$RV\$GLR\$KLDL\$D\$LTNRNQFVD\$A\$FG\$SLMDDI\$G\$VQTVVNAVS	729
Db	612	ANVFEDLENILREYNSYK---SALYTI\$EAKIATNT\$PYVNG\$INSF\$LQ\$G\$G\$AIGT\$G\$V\$IS	668
Qy	730	GVAITL\$FSIVTGFIN\$IKNP\$FGG\$MLMI\$V\$IG\$V\$LP\$AI\$YFL\$TKT\$KI\$YET\$API\$KMI\$YP---	786
Db	669	VTACALGDI\$VG\$V\$V\$F\$LN\$K\$NP\$FGG\$GL\$ML\$AI\$V\$V\$V\$II\$V\$V\$F\$VR\$QR\$H\$V\$LSK\$PI\$DM\$M\$FP\$AT	728
Qy	787	-----ETDKL\$KERECK\$SIAPI\$SEEE	807
Db	729	NPVTTV\$SV\$TGT\$TVV\$V\$T\$P\$SV\$K\$DV\$GGT\$SVA-V\$SEKE	763

```

RESULT 38
US-09-612-204B-2
; Sequence 2, Application US/09612204B
; Patent No. 6461811
; GENERAL INFORMATION:
; APPLICANT: Patience, Clive
; TITLE OF INVENTION: Swine Gamma Herpesvirus DNA and Methods of Use
; FILE REFERENCE: 61750-299
; CURRENT APPLICATION NUMBER: US/09/612,204B
; CURRENT FILING DATE: 2001-08-13
; PRIOR APPLICATION NUMBER: U.S. 60/142,736
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: U.S. 60/168,532
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Deduced amino
; OTHER INFORMATION: acid sequence derived from the first open reading
; OTHER INFORMATION: frame of the DNA of SED ID NO:1
US-09-612-204B-2

```

	Query Match	22.0%; Score 1002; DB 4;	Length 195;
	Best Local Similarity	100.0%; Pred. No. 2.3e-70;	
	Matches 195; Conservative 0;	Mismatches 0; Indels 0;	Gaps 0
Qy	484	NLRIRINNILEDISKAWCREQHRAALVWNELSKINPTSVMSMIYNRPVSAKRIGDVISVS	543
Dd	1	NLRIRINNILEDISKAWCREQHRAALVWNELSKINPTSVMSMIYNRPVSAKRIGDVISVS	60
Qy	544	NCIVDQTSLHKLRLLSASDEKCFSRPVPVTFKFMDNSTIYKGQLGVNNEILLTTTTL	603
Dd	61	NCIVDQTSLHKLRLLSASDEKCFSRPVPVTFKFMDNSTIYKGQLGVNNEILLTTTTL	120
Qy	604	ETCOENTEYFQAOKDMVIYKNYBHLKTVPULSSITTLDTFFIALNFTLLENVDKFKIELYT	663

```

Db      121  ETCQENTYYFQAKDMDYIKNVEHLKTVPLSSITTTDLTFIALNFTLLENVDFKVIELYT 180
Qy      664  RDEKRLSNVFDIETM 678
          |||||
Db      181  RDEKRLSNVFDIETM 195

RESULT 39
US-08-220-151-2
; Sequence 2, Application US/08220151
; Patent No. 5529780
; GENERAL INFORMATION:
;   APPLICANT: Paoletti, Enzo
;   APPLICANT: Limbach, Keith J.
;   TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
;   TITLE OF INVENTION: CANINE HERPESVIRUS gB, gC AND gD AND USES THEREFOR
;   NUMBER OF SEQUENCES: 91
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Curtis, Morris & Safford
;   STREET: 530 Fifth Avenue
;   CITY: New York
;   STATE: NY
;   COUNTRY: USA
;   ZIP: 10036
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/220,151
;   FILING DATE: 30-MAR-1994
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Frommer, William S.
;     REGISTRATION NUMBER: 25,506
;     REFERENCE/DOCKET NUMBER: 454310-2540
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (212) 840-3333
;     TELEFAX: (212) 840-0712
;     TELEX: 425066 CURTMS
;   INFORMATION FOR SEQ ID NO: 2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 879 amino acids
;       TYPE: amino acid
;       STRANDEDNESS: single
;       TOPOLOGY: linear
;     MOLECULE TYPE: peptide
;     FRAGMENT TYPE: internal
;   US-08-220-151-2

```

[illegible]

Db 233 YDSFAISTGDIHMSPPFGLRDLGAHTYBISYSTDRFQOIEGYPIIDLDTLRLQLGAPVSRN 292
Qy 314 FLDRBEYTLGWEKHLKMSYCPPLTLKAFNGIQTGSHGSHFVANDITASF--TTSKED 371
Db 293 FLTQHTVAVWVWPKIREVCTTLAKWREIDBIIRDEYKGSYRETAKSISATFISDTTQFD 352
Qy 372 MKEFTNTYHCLNEBEIKAEIEKY-AKVNSTHSGYDLKPKYKTDGGLYLVMOPLI----- 424
Db 353 IDRVKLS-DCAKREAIETADKIYKKYKNTHTIQTGELETYLAGGFIAPRPMSNELAK 411
Qy 425 -----QNRLDANKLN-----NETYSRRRQAEISTTDPMMETGNGAGG 465
Db 412 LYINELVRSNRTVDLKSLLNPSVRGARKRRSVEENKSKRNIEGGIENVNST----- 465
Qy 466 EYSSENSITVAQVAYADNLRIRINNILEDLSKACREQHRALVWNELSKINPTSVMSM 525
Db 466 IIKTSSVHFAMLOFAYDHIQSHVNEMLSRATACNQLKQKERTLWNEVWKLNPTSVASV 525
Qy 526 IYNRPVSAKRIGDVISVSNICIVDQTSVSLHKSURLLSASDEKCFSPRPVTFKFMNDSTI 585
Db 526 AMDQVSARMGLDVLAVTQCVNISGSSVFQNSMRVL-GSTTTCYSRPLSFKALENSTN 584
Qy 586 Y-KGQLGVNNEILLTITTYLETQCENTEYFQAKTDMYIKNYEHLKTVPLSSITLDTFI 644
Db 585 YIEQQLGENNELVERKLEIEPCTANHKKRYKFGADYVVFENYAVRVKPLNEIEMISAYV 644
Qy 645 ALNFTLLENVDKVIETYTRDEKRLSNVFDIETMFREYNYYAQRVSGRLKDLDLSTNRN 704
Db 645 DLNITLLEDFLEFVYTRAELEDTGLLDYSIQRRNQLHALKFYDI--DSVVKVDNNV 702
Qy 705 QFVDAFGSLMDDLGAVGQTVVNAVSGVATLTFSSIVTGFINFKNPFGGMLMIIVIGVLF 764
Db 703 VIMRGIANFQGLGDVGAGFKVVLGAANAVIATVSGVSFLNPPFALAVGLLILAGLF 762
Qy 765 AIYFLTKTKIYETAPIKMIYPEIDK-LKE-----REGKSEIAPISEBELER--- 810
Db 763 AAPLAYRVYVKLSNPMKALYPVTTKNLKESVKNNGSNGNSDGEENDNDIEEKLOAKE 822

RESULT 40
US-08-220-151-3
Sequence 3, Application US/08220151
Patent No. 5529780
GENERAL INFORMATION:
APPLICANT: Paolletti, Enzo
APPLICANT: Limbach, Keith J.
TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
TITLE OF INVENTION: CANINE HERPESVIRUS 9B, 9C AND 9D AND USES THEREFOR
NUMBER OF SEQUENCES: 91
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford
STREET: 530 Fifth Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/220,151
FILING DATE: 30-MAR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Frommer, William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2540

TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 840-3333
TELEFAX: (212) 840-0712
TELEX: 425066 CURTMS
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 879 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: N-terminal
US-08-220-151-3

Query Match 19.8%; Score 901; DB 1; Length 879;
Best Local Similarity 28.1%; Pred. No. 1.7e-61;
Matches 252; Conservative 168; Mismatches 402; Indels 76; Gaps 21;

Qy 17 LYQVALYSLSIABTGVTSPPNTATWSTESPLTGHYGTHTDSSHGCGNENNRDSEEQNKNI 76
Db 1 MFSLYIYIFIIYLLIICDPTTPE-STINPLNHNHNLSTPKPTSDDIETILRESQIE----- 55
Qy 77 YGSPSTFPYRVCSAGSGVDVFRQTDHVCPCDAS-DMVHSEGLLIYKQNIIPPMFVRKY 135
Db 56 --SDDTSTFVMCPSPSGSTLVRLPPRACPNYKLGKNTFTGIAVIFKENISPYKFRANIY 113
Qy 136 RKVVTSTTVNGIYSDSIINQHTFYKSI-EPWETEKMDTIYQCFNSRLRNTTGNLLTYVD 194
Db 114 YKNIITITVMSGSTYAVITNRYTDRVPIGVPEITELIDRRGMCLSKADYIRNNYEFTAFD 173
Qy 195 RDDINNTVFLQPDVGTVPDKRYGSOPELYLEPGWFGSYRRRTTVNCELMDMFARSNPP 254
Db 174 KUEDPREVHLKSKFNTPSRGHVTNDYTKIGGS-GFYHSGTSVNCIVEEDARSVP 232
Qy 255 FDFVTFATGDTVEMSPFWSGEDD-HENKHEKPMFVSVINNYKVVDYQNRGTVPGLKTRI 313
Db 233 YDSFAISTGDIHMSPPFGLRDLGAHTYBISYSTDRFQOIEGYPIIDLDTLRLQLGAPVSRN 292
Qy 314 FLDRBEYTLGWEKHLKMSYCPPLTLKAFNGIQTGSHGSHFVANDITASF--TTSKED 371
Db 293 FLTQHTVAVWVWPKIREVCTTLAKWREIDBIIRDEYKGSYRETAKSISATFISDTTQFD 352
Qy 372 MKEFTNTYHCLNEBEIKAEIEKY-AKVNSTHSGYDLKPKYKTDGGLYLVMOPLI----- 424
Db 353 IDRVKLS-DCAKREAIETADKIYKKYKNTHTIQTGELETYLAGGFIAPRPMSNELAK 411
Qy 425 -----QNRLDANKLN-----NETYSRRRQAEISTTDPMMETGNGAGG 465
Db 412 LYINELVRSNRTVDLKSLLNPSVRGARKRRSVEENKSKRNIEGGIENVNST----- 465
Qy 466 EYSSENSITVAQVAYADNLRIRINNILEDLSKACREQHRALVWNELSKINPTSVMSM 525
Db 466 IIKTSSVHFAMLOFAYDHIQSHVNEMLSRATACNQLKQKERTLWNEVWKLNPTSVASV 525
Qy 526 IYNRPVSAKRIGDVISVSNICIVDQTSVSLHKSURLLSASDEKCFSPRPVTFKFMNDSTI 585
Db 526 AMDQVSARMGLDVLAVTQCVNISGSSVFQNSMRVL-GSTTTCYSRPLSFKALENSTN 584
Qy 586 Y-KGQLGVNNEILLTITTYLETQCENTEYFQAKTDMYIKNYEHLKTVPLSSITLDTFI 644
Db 585 YIEQQLGENNELVERKLEIEPCTANHKKRYKFGADYVVFENYAVRVKPLNEIEMISAYV 644
Qy 645 ALNFTLLENVDKVIETYTRDEKRLSNVFDIETMFREYNYYAQRVSGRLKDLDLSTNRN 704
Db 645 DLNITLLEDFLEFVYTRAELEDTGLLDYSIQRRNQLHALKFYDI--DSVVKVDNNV 702
Qy 705 QFVDAFGSLMDDLGAVGQTVVNAVSGVATLTFSSIVTGFINFKNPFGGMLMIIVIGVLF 764
Db 703 VIMRGIANFQGLGDVGAGFKVVLGAANAVIATVSGVSFLNPPFALAVGLLILAGLF 762
Qy 765 AIYFLTKTKIYETAPIKMIYPEIDK-LKE-----REGKSEIAPISEBELER--- 810
Db 763 AAPLAYRVYVKLSNPMKALYPVTTKNLKESVKNNGSNGNSDGEENDNDIEEKLOAKE 822

QY 811 -----IVLAMIHQONSHMETKTRKDPKDSILTRAQNMRLKRG--YSNLKNAESVE 860
 Db 823 MIKWSLVSAM---EQEHRKAIKNSGPA-LLASHITNLSLKHGPKYKRLKVNENE 876
 RESULT 41
 US-08-413-118-2
 ; Sequence 2, Application US/08413118
 ; Patent No. 5688920
 ; GENERAL INFORMATION:
 ; APPLICANT: PAOLETTI, ENZO
 ; APPLICANT: LIMBACH, KEITH J.
 ; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
 ; TITLE OF INVENTION: CANINE HERPESVIRUS 9B, 9C, AND 9D AND USES THEREFOR
 ; NUMBER OF SEQUENCES: 128
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
 ; STREET: 530 FIFTH AVENUE, 25TH FLOOR
 ; CITY: NEW YORK
 ; STATE: NEW YORK
 ; COUNTRY: UNITED STATES OF AMERICA
 ; ZIP: 10036
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/413,118
 ; FILING DATE: 29-MAR-1995
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/220,151
 ; FILING DATE: 30-MAR-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: FROMMER, WILLIAM S.
 ; REGISTRATION NUMBER: 25,506
 ; REFERENCE/DOCKET NUMBER: 454310-2670
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 840-3333
 ; TELEFAX: (212) 840-0712
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 879 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; FRAGMENT TYPE: internal
 US-08-413-118-2
 Query Match 19.8%; Score 901; DB 1; Length 879;
 Best Local Similarity 28.1%; Pred. No. 1.7e-61;
 Matches 252; Conservative 168; Mismatches 402; Indels 76; Gaps 21;
 QY 17 LYQALYSLSIAETGVSPNTATWSTESPLTGHVGHDSHGGRGNENRDSBQNKNI 76
 Db 1 MFSLYLIFFIYLIICDPTPE-STINPLNHNLSLTKPTSDDIRILRESQIE---- 55
 QY 77 YGSPSTFFPYRVGSGVDVFRFQDHYVCPDAS-DMVHSEGIILLYKQNIIPFMRVRKY 135
 Db 56 --SDTSTFYMCPPPSGSLVRLPRACPNVYKLGKFTGIAVIFKENISFYKFKANIY 113
 QY 136 RKVTTSTVYNGIYSDSTINQHTFYKSI-EPWETEKMDTIYQCFNSRLNTGGNLLTYVD 194
 Db 114 YKNIIITVWSGSTYAVITNRYTRDVPVIGVPEITELIDRRGCLSKADYIRNNVEFTAFD 173
 QY 195 RDDINWTVFLOPDVGTDPVKYGSQPELYLEPGHFWGYSRRRTVNCELMDMARSNPP 254
 Db 174 KDDEPREVHLKPSKFNTPGSGWHTVNDTYTKIGGS-GFYHSGTSVNCIVEVDARSYP 232
 QY 255 FDFVFTATGDTVMSPFWSGDD-HENKMKHPFVSVINNVKVVYDQNRGTVPGLKTRI 313

Db 233 YDSPAISTGDIHHSPPFGLRGAHTVEISYSTDRFQOIEGYPIIDLTQLGAPVSRN 292
 QY 314 FLDRREYTLSEKHLKWSYCPILWLKAFYNGIQTEHSGSYHFVANDITASF--TTSKED 371
 Db 293 FLTTQHTVAVNWVPKIREVCTLAKWREIDRIIRDEYKGSYRFTAKSISATFISDTQFD 352
 QY 372 MKEFNNTVHCLNEBEIKABIEKKY-AKVNSTHSHKYGDLKYFKTGGGLYLVWQPLI----- 424
 Db 353 IDRVKLS-DCAKRAIEAIDKIYKKYKNTHTIQGELETYLRGGFFIAFRPMISNEAK 411
 QY 425 -----QNRLLDANKLN-----NETYSRRSRQAESTTDDPMEMTNGAGG 465
 Db 412 LYINELVRSNRTVDLKSLLNPSVRGGARKRSVEENKSKRNIEGGIENVNST----- 465
 QY 466 EYSSENSITVAQVAYDNLRIRINNIILEDLSKAWCREQHRRAALVWNLKINPTSVMSM 525
 Db 466 IIKTSSVHFAMLOFAYDHIQSHVNEMLSRITATAMCNLQNKERTLWNEVMKLNPTSVASV 525
 QY 526 IYNRPVSAKRIGDVISVNSCIWVDQTSVSLHKSRLLSASDEKCFSPRPVTFKPMNDSTI 585
 Db 526 AMDQVSARMGLGDVLAVTQCVNIGSSVFIQNSMRVL-GSTTTCYSRRLISFKALENSTN 584
 QY 586 Y-KGQLGVNNEILLTITTYLETQCENTEYFOAKTDMYIKNYEHLKTVPLSSITTLDTFI 644
 Db 585 YIEQQLGENNELVERKLEIETCTANHKRYFKGADYVYFENYAVYRKYPLNEIEMISAVV 644
 QY 645 ALNFTILENVDFKVIELYTRDEKRLSNVFDIETMFREYNYAQRVSGURKOLLDLSTVRN 704
 Db 645 DLNITLDEDFLEVYTRAELEDTGLLDYSEIQRNQLHALKFYDI--DSVVKVDNVV 702
 QY 705 QFVDAFGLMDDLGAVGTVVNAVSGVATLPSSTVTGTFINFIKNPFGGMLIIVIGVLF 764
 Db 703 VMRGIANFFOGLGDVGAGFGKVLGAANAVIATVSGVSSFLNNPFGALAVGLLLLAGLF 762
 QY 765 AIYFLTKTKIYETAPIKMIYPEIDK-LKE-----REGKSEIAPISSEELER--- 810
 Db 763 AAFLAYRVVSKLNSNPKALYPTVTKNLKSQVKNNGNSGNSDGEENDNDIDEKLOOAKE 822
 QY 811 -----IVLAMIHQONSHMETKTRKDPKDSILTRAQNMRLKRG--YSNLKNAESVE 860
 Db 823 MIKWSLVSAM---EQEHRKAIKNSGPA-LLASHITNLSLKHGPKYKRLKVNENE 876
 RESULT 42
 US-08-413-118-3
 ; Sequence 3, Application US/08413118
 ; Patent No. 5688920
 ; GENERAL INFORMATION:
 ; APPLICANT: PAOLETTI, ENZO
 ; APPLICANT: LIMBACH, KEITH J.
 ; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
 ; TITLE OF INVENTION: CANINE HERPESVIRUS 9B, 9C, AND 9D AND USES THEREFOR
 ; NUMBER OF SEQUENCES: 128
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
 ; STREET: 530 FIFTH AVENUE, 25TH FLOOR
 ; CITY: NEW YORK
 ; STATE: NEW YORK
 ; COUNTRY: UNITED STATES OF AMERICA
 ; ZIP: 10036
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/413,118
 ; FILING DATE: 29-MAR-1995
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/220,151
 ; FILING DATE: 30-MAR-1994


```
Db 174 KDEPREVHLKPSKFNTPGSRGWHTVNDYTTKIGGS-GFYHSGTSVNCIVEVDARSVYP 232
Qy 255 PFFVATGDTVMGSPFWSGEDD-HENKMHKPMFVSVNNYKVVDYQNRGTVPLGKTRI 313
Db 233 YDSFAISTGDIITHMFPFGLRGCAHTEVYSYSTRFQOIEGYPIDLTRLQLGAPVSRN 292
Qy 314 FLDBREYTLWSBKHLKNMSYCPPLTLWKAIFYNGIOTHSYHFAVANDITASF--TTSKED 371
Db 293 FLTQHVTVAMNWVPKIREVCTLAKWREIDEIIRDEYKGSYRFTAKSISATISDTTQFD 352
Qy 372 MKEFNNTYHCLNEBEIKAEIKKY-AKVNSTHSGYDLYKFKTDGGLYLWQPLI----- 424
Db 353 IDRVKLS-DCAKREAIERDAIKYKKYKNTHTIQTGELETYLARGGFIAPFPMISNELAK 411
Qy 425 -----QNRLLDAKNKLN-----NETYSRRSRQAESTTDPMMEMTNGAGG 465
Db 412 LYINELVSRNRTVDLKSLLNPSVRGARGARRSVEENKSRKNIEGGIENVNST----- 465
Qy 466 EYSENSITVAQVQYAYDNLRIINNILEDLSKAWCREQHRAALVWNLKINPTSVMSM 525
Db 466 IIKTTSSVHFAMLQFAYDHIQSHVNEMLSRATAMCNLQNKERTLWNEVMKLNPTSVASV 525
Qy 526 IYNRPVSAKRIGDVISVNCIVVDQTSVLSLHLSASDEKCFSPRPVTFKFMNDSTI 585
Db 526 AMDQVSARMGLDVLAVTQCUNISGSSVFIQNSMRVL-GSTTTCYSRPLISFKALENSTN 584
Qy 586 Y-KGQLGVNNEILLTTTLETQCENTEYVFOAKTDMYIKNYEHLKTVPLSITLDTFI 644
Db 586 Y-KGQLGVNNEILLTTTLETQCENTEYVFOAKTDMYIKNYEHLKTVPLSITLDTFI 644
Qy 645 ALNFTLLENVDFKVIELYTRDEKRLSNVDFIETMFREYNYAQRVSGLRKOLLDLSTNRN 704
Db 645 DLNITLLEDFLEPLVYTRAELEDGLDYSEIQRRNLHALKFYDI--DSVVKVDNNV 702
Qy 705 QFVDAFGLSMDLGLAVGTQVNVAVSGVATLFSISIVTGFINFKNPFGLMGLIIVIGVLF 764
Db 703 VMIRGIANFFQGLGVDGAGFGKVLGAANAVIATVSGVSSFLNPPFGALAVGLLILAGLF 762
Qy 765 AIYFLTKTKIYETAPIKMIYPEIDK-LKE-----REGKSEIAPISEEELER--- 810
Db 763 AAFLAYRVSKLSPMKALYPVTTKNLKESVNGNSGNSDGEENDNDIDEKLIQOAKE 822
Qy 811 -----IVLAMHIHQNSHMETKTRKDPKDSILTTRAQNMRLKRSR--YSNLKNAESVE 860
Db 823 MIKWSLVSAM--EQOEHAIKKNSGPA-LLASHITNLSLKHRCPKYKRLKVNENE 876

RESULT 44
US-08-473-446-2
; Sequence 2, Application US/08473446
; Patent No. 6017542
; GENERAL INFORMATION:
; APPLICANT: PAOLETTI, ENZO
; APPLICANT: LIMBACH, KEITH J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; TITLE OF INVENTION: CANINE HERPESVIRUS gb, gc, and gd AND USES THEREFOR
; NUMBER OF SEQUENCES: 128
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
; STREET: 530 FIFTH AVENUE, 25TH FLOOR
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/473,446
; FILING DATE:
; CLASSIFICATION:
```

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/413,118
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: FROMMER, WILLIAM S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2670
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 879 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; US-08-473-446-2

Query Match 19.8%; Score 901; DB 3; Length 879;
Best Local Similarity 28.1%; Pred. No. 1.7e-61;
Matches 252; Conservative 168; Mismatches 402; Indels 76; Gaps 21;

Qy 17 LYQVALYSLSIAETGVTSPPTATWSTESPLTGHYTHDSHGGERGNENRDSBEQNKNI 76
Db 1 MFSLYLFIFFIYLIICDPTPE-STINPLNHNLSLTPPKPTSDDIREILRESQIE----- 55
Qy 77 YGSPSTPYRVCSAGVDVRFOTDHVCPDAS-DMVHSEGLIYKNIIPFPRVKRY 135
Db 56 --SDTSTFYMCPPSPGSTLVRLEPPRACPNYKLGKNTGEGIAVIFKENISPYKFKANIY 113
Qy 136 RKVVTTVYNGIYSDTSITNQHTFYKSI-EPWETEKMDTIYQCNSLRNLTTGGNLLTYVD 194
Db 114 YKNIITTVSGSYAVITNRYTRDVRPIGVPEITELIDRRGMCLSKADYIRNNVEFTAPD 173
Qy 195 RDDINMTVFQPDGVPTDVKRYGQPELYLEPCWFGWSYRRRTTVNCELMDMFARSNPP 254
Db 174 KDEPREVHLKPSKFNTPGSRGWHTVNDYTTKIGGS-GFYHSGTSVNCIVEVDARSVYP 232
Qy 255 PFFVATGDTVMGSPFWSGEDD-HENKMHKPMFVSVNNYKVVDYQNRGTVPLGKTRI 313
Db 233 YDSFAISTGDIITHMFPFGLRGCAHTEVYSYSTRFQOIEGYPIDLTRLQLGAPVSRN 292
Qy 314 FLDBREYTLWSBKHLKNMSYCPPLTLWKAIFYNGIOTHSYHFAVANDITASF--TTSKED 371
Db 293 FLTQHVTVAMNWVPKIREVCTLAKWREIDEIIRDEYKGSYRFTAKSISATISDTTQFD 352
Qy 372 MKEFNNTYHCLNEBEIKAEIKKY-AKVNSTHSGYDLYKFKTDGGLYLWQPLI----- 424
Db 353 IDRVKLS-DCAKREAIERDAIKYKKYKNTHTIQTGELETYLARGGFIAPFPMISNELAK 411
Qy 425 -----QNRLLDAKNKLN-----NETYSRRSRQAESTTDPMMEMTNGAGG 465
Db 412 LYINELVSRNRTVDLKSLLNPSVRGARGARRSVEENKSRKNIEGGIENVNST----- 465
Qy 466 EYSENSITVAQVQYAYDNLRIINNILEDLSKAWCREQHRAALVWNLKINPTSVMSM 525
Db 466 IIKTTSSVHFAMLQFAYDHIQSHVNEMLSRATAMCNLQNKERTLWNEVMKLNPTSVASV 525
Qy 526 IYNRPVSAKRIGDVISVNCIVVDQTSVLSLHLSASDEKCFSPRPVTFKFMNDSTI 585
Db 526 AMDQVSARMGLDVLAVTQCUNISGSSVFIQNSMRVL-GSTTTCYSRPLISFKALENSTN 584
Qy 586 Y-KGQLGVNNEILLTTTLETQCENTEYVFOAKTDMYIKNYEHLKTVPLSITLDTFI 644
Db 586 Y-KGQLGVNNEILLTTTLETQCENTEYVFOAKTDMYIKNYEHLKTVPLSITLDTFI 644
Qy 645 ALNFTLLENVDFKVIELYTRDEKRLSNVDFIETMFREYNYAQRVSGLRKOLLDLSTNRN 704
Db 645 DLNITLLEDFLEPLVYTRAELEDGLDYSEIQRRNLHALKFYDI--DSVVKVDNNV 702
Qy 705 QFVDAFGLSMDLGLAVGTQVNVAVSGVATLFSISIVTGFINFKNPFGLMGLIIVIGVLF 764
Db 703 VMIRGIANFFQGLGVDGAGFGKVLGAANAVIATVSGVSSFLNPPFGALAVGLLILAGLF 762
Qy 765 AIYFLTKTKIYETAPIKMIYPEIDK-LKE-----REGKSEIAPISEEELER--- 810
Db 763 AAFLAYRVSKLSPMKALYPVTTKNLKESVNGNSGNSDGEENDNDIDEKLIQOAKE 822
Qy 811 -----IVLAMHIHQNSHMETKTRKDPKDSILTTRAQNMRLKRSR--YSNLKNAESVE 860
Db 823 MIKWSLVSAM--EQOEHAIKKNSGPA-LLASHITNLSLKHRCPKYKRLKVNENE 876
```

Db 703 VMRGIANFFQGLGVDVAGGKVVVLGAANAVIATVSGVSSFLNPPFALAVGLLILAGLF 762
Qy 765 AIYFLTKTKIYETAPIKMIYPEIDK-LKE-----REGKSEIAPISEBELER--- 810
Db 763 AAFLAYRYVSKLSPNPKALYPVTTKNLKEGVKNGSGNNDGDENDNDIDEEKLOAKE 822
Qy 811 -----IVLAMHIHQONSHMETKTRKDPKDSILTRAQNMMLRKRSQ--YSLNKAESVE 860
Db 823 MIKYSLSVSAM---EQOEHAIKKNSGPA-LLASHITNLSLKHGPKYKRLKNVNE 876

RESULT 45
US-08-473-446-3
; Sequence 3, Application US/08473446
; Patent No. 6017542
; GENERAL INFORMATION:
; APPLICANT: PAOLETTI, ENZO
; APPLICANT: LIMBACH, KEITH J.
; TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES OF
; TITLE OF INVENTION: CANINE HERPESVIRUS 9B, 9C, AND 9D AND USES THEREFOR
; NUMBER OF SEQUENCES: 128
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
; STREET: 530 FIFTH AVENUE, 25TH FLOOR
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/473,446
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/413,118
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: FROMMER, WILLIAM S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454310-2670
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 879 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
US-08-473-446-3

Query Match 19.8%; Score 901; DB 3; Length 879;
Best Local Similarity 28.1%; Pred. No. 1.7e-61;
Matches 252; Conservative 168; Mismatches 402; Indels 76; Gaps 21;

Qy 17 LYQVALYSLIAETGVTSPNTATWSTESPLTGHYTHDSSHGSGNGNENRDSEBQNKNI 76
Db 1 MFSLYLFIIFIYTLITCDTPPE-STINPLNHNLSLTPKTSDDIRILRESQIE--- 55
Qy 77 YGSPSTPPYVCSGASGVDFRFTDHWCPDAS-DMVHSEGLILYKONIIPEMFVRKY 135
Db 56 --SDDSTFFWCPPPGSGLVRLEPPRACPNYKLGKNFTGEGIAVFKENISPYKFKANIY 113
Qy 136 RKVVTSTVNGIYSDITNQHTFYKSI-EPWETEKMDTIYQCFSNLSRLNTGGNLLTYVD 194
Db 114 YKQIITTVWSGTYAVITNRYTRDVPVIGVPEITELIDRGMCLSKADYIRNNVEFTAFD 173

Search completed: October 28, 2003, 15:30:06
Job time : 42 secs

Qy 195 RODINMTVFLQPDGVGTDPVKRYGSOPELYLEPGFWGYSRRRTTYNCMLMDMFARSNPP 254
Db 174 KOEDPREVHLKSKFNTPGSRGHWHTVNDYTKLIGS-GFYHSGTSSVNCIVEEDARSVP 232
Qy 255 FPFVFTATGDTVEMSPFWSGEDD-HENKHEKPMFWFVSVINNYKVVDYQNRGTVPGLGTRI 313
Db 233 YDSFALSTGDIIMSPFFGLRDCGATEYISYSTDRFQQIEGYYPIDLDLTLQLGAPVSRN 292
Qy 314 FLDBREYTTLSWEKHLKNMSYCPITLWKAPYNGIQTEHSGSYHFAVNDITASF--TTSKED 371
Db 293 FLTTQHTVTAVNWPVKIREVCTLAKWREIDEIIRDEYKGSYRFTAKSISATFISDTQPD 352
Qy 372 MKEFNTHYHCLNEEIKAEIEKKY-AKVNSHYSKYGLKVKFKTDGGLYLVWQPLI----- 424
Db 353 IDRVKLS-DCAKREATEADKIYKKYKNTHTQTGELETYLARGGFIIFAPRPMISLAK 411
Qy 425 -----QNRLLDAKNKLN-----NETYSRRSRQAESTTDPMMMTGNGAGG 465
Db 412 LYINELVRSNRTVDLKSLLNPSVRGARGKRSVEENKRSKRNIEGGIENVNST----- 465
Qy 466 EYSSSENSITVAQVQYAYDNLIRINNIBDLKAWCREQHRALVWNELSKINPTSVMSM 525
Db 466 IIKTSSVHFAMLQFAYDHIQSHVNEMLSRIATAWCNQLONKERTLWNEVMKLNPTSVASV 525
Qy 526 IYNRPVSAKRIGDVISVSNICVVDDQTSVSLHLSLRLSASDEKCFSRPPVTFKFMNDSTI 585
Db 526 AMDQVRSARMLGDLAVTQCVNISGSSVFIONSMRVL-GSTTTCYSRPLISFKALENSTN 584
Qy 586 Y-KGQLGVNNEILLTTTLETQENTYFYFOAKTDMYIYKNEYHLKTVPLSSITTLDTFI 644
Db 585 YIEGQLGENNELLVERKLEPCTANHRYKFKGADYVYFENYAVYKVPLENEIMISAYV 644
Qy 645 ALNFTLLENVDVKVIELYTRDEKRLSNVFDIETMPREYNYAQRVSGGLRKDLDLSTNRN 704
Db 645 DLNITLLEDEFLPLEVYTRAELEDGLDYSIEIQRNQLHALKPYDI--DSVVKVDNVV 702
Qy 705 QFVDAFGSLMDDLGAQGVTVNNAVSGVATLPSISIVTGFNFINKNPGFGLMIITVIGVLF 764
Db 703 VMRGIANFFQGLGVDVAGGKVVVLGAANAVIATVSGVSSFLNPPFALAVGLLILAGLF 762
Qy 765 AIYFLTKTKIYETAPIKMIYPEIDK-LKE-----REGKSEIAPISEBELER--- 810
Db 763 AAFLAYRYVSKLSPNPKALYPVTTKNLKEGVKNGSGNNDGDENDNDIDEEKLOAKE 822
Qy 811 -----IVLAMHIHQONSHMETKTRKDPKDSILTRAQNMMLRKRSQ--YSLNKAESVE 860
Db 823 MIKYSLSVSAM---EQOEHAIKKNSGPA-LLASHITNLSLKHGPKYKRLKNVNE 876

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 15:27:11 ; Search time 86 Seconds
(without alignments)
1596.494 Million cell updates/sec

Title: US-10-055-364-24

Perfect score: 4547

Sequence: 1 MAGSLKRLGSLVLAFLWLYQV.....KRGYSNLKNAESVEMLNTL 865

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_19Jun03.*

- 1: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
- 24: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4547	100.0	865	22 AAB31679	Amino acid sequenc
2	1937	42.6	845	18 AAW26475	KSHV glycoprotein
3	1927.5	42.4	829	21 AAB53128	Macaca mulatta rha
4	1840	40.5	856	22 AAUI4094	Peptide sequence f
5	1798.5	39.6	927	8 AAP70768	Epstein-Barr virus
6	1128	24.8	907	15 AAP63462	Human cytomegalovi
7	1123	24.7	907	18 AAW27274	Human cytomegalovi
8	1121.5	24.7	906	8 AAP70289	Sequence encoded b
9	1114	24.5	906	10 AAP91047	gb envelope protei

10	1002	22.0	195	22	AAB31678	A viral glycoprote
11	901	19.8	879	16	AAR77024	Canine herpesvirus
12	876.5	19.3	865	14	AAR30169	Marek's Disease Vi
13	851.5	18.7	933	13	AAR27807	Bovine herpes viru
14	851.5	18.7	933	13	AAR41343	Bovine herpesvirus
15	847.5	18.6	904	12	AAR41465	HSVgB polypeptide.
16	845.5	18.6	903	8	AAP71135	Herpes Simplex Vir
17	844	18.6	904	23	AAE17812	Herpes simplex vir
18	843.5	18.6	903	8	AAP70426	Recombinant herpes
19	841	18.5	904	17	AAW00375	HSV-1 glycoprotein
20	841	18.5	904	22	AAB74441	Herpes simplex vir
21	839	18.5	854	19	AAW72113	HSV-2 strain SB5 C
22	839	18.5	904	14	AAW41778	Glycoprotein B (gB
23	839	18.5	904	17	AAW00376	HSV-2 glycoprotein
24	839	18.5	904	19	AAW72193	HSV-2 strain SB5 C
25	839	18.5	904	22	AAB74442	Herpes simplex vir
26	838.5	18.4	903	7	AAP60244	Herpes simplex vir
27	838	18.4	885	17	AAK92747	SAB virus gb glyco
28	838	18.4	904	18	AAW34552	Herpes simplex vir
29	838	18.4	907	8	AAP71136	Herpes Simplex Vir
30	836.5	18.4	904	14	AAW41779	Glycoprotein B (gB
31	836	18.4	795	19	AAW72062	HSV-2 strain SB5 C
32	833	18.3	904	9	AAP80914	Sequence of Herpes
33	832	18.3	868	13	AAK22615	Varicella-zoster v
34	829.5	18.2	904	18	AAW34553	Herpes simplex vir
35	823.5	18.1	905	9	AAP80915	Sequence of Herpes
36	821	18.1	873	16	AAW79459	Infectious Laryngo
37	819.5	18.0	973	8	AAP70769	Glycoprotein B of
38	818.5	18.0	891	17	AAK92746	B virus gb glycopr
39	818.5	18.0	891	19	AAW70293	Simian herpesvirus
40	811	17.8	854	8	AAP70347	Varicella-zoster v
41	808.5	17.8	903	6	AAP50312	Herpes simplex vir
42	789.5	17.4	943	21	AAV32470	DNA encoding felin
43	789.5	17.4	943	22	AAW50113	Feline herpesvirus
44	777.5	17.1	913	19	AAW68404	Aujeszky's disease
45	757	16.6	1086	18	AAW36051	Hybrid Marek's dis

ALIGNMENTS

RESULT 1

AAB31679
ID AAB31679 standard; Protein; 865 AA.

XX AAB31679;

AC AAB31679;

XX 30-APR-2001 (first entry)

DT Amino acid sequence of a viral glycoprotein B envelope protein.

DE Glycoprotein B; envelope protein; vaccine; passive immunity; pig.

XX Swine gamma herpesvirus.

XX WO200103733-A1.

PD 18-JAN-2001.

XX 07-JUL-2000; 2000WO-US18647.

XX 08-JUL-1999; 99US-0142736.

PR 02-DEC-1999; 99US-0168532.

XX (BIOT-) BIOTRANSPLANT INC.

XX Patience C;

XX MPI; 2001-103033/11.

DR N-PSDB; AAP25199.

XX New polynucleotide encoding a polypeptide corresponding to porcine

PT gamma-herpesvirus glycoprotein B (pGHV-gpB) for immunizing a pig

PT against a porcine gamma-herpesvirus -
PS Claim 9; Fig 3; 82pp; English.
XX

CC The present sequence represents a swine gamma herpesvirus glycoprotein
CC B envelope protein. Probes and primers derived from DNA encoding the
CC glycoprotein B envelope protein are used for detecting the presence
CC of swine gamma herpesvirus. Antibodies directed against the envelope
CC protein are also useful for detecting presence of the virus. The
CC envelope protein can be used as a vaccines for creating passive
CC immunity in a pig.
XX
SQ Sequence 865 AA;

Query Match 100.0%; Score 4547; DB 22; Length 865;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 865; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGSLKLRGSLVLAALWLYQVALYSLTAETGVTSPPTATWTSTESPLTGHYTHDSSHGE 60
DB 1 MAGSLKLRGSLVLAALWLYQVALYSLTAETGVTSPPTATWTSTESPLTGHYTHDSSHGE 60
QY 61 RGNENRDSBEQNKNIYGSPTFPYRVCSAGSGVDVFRFQTDHVCPCDASDMVHSEGIILLI 120
DB 61 RGNENRDSBEQNKNIYGSPTFPYRVCSAGSGVDVFRFQTDHVCPCDASDMVHSEGIILLI 120
QY 121 YKQNIIPFMFRVRKYRKVVTSTVYNGIYSDSITNQHTFYKSIETPWTETKMDTIYQCFSN 180
DB 121 YKQNIIPFMFRVRKYRKVVTSTVYNGIYSDSITNQHTFYKSIETPWTETKMDTIYQCFSN 180
QY 181 LRLNTGGLTYVDRDINMTVFLQPDVGTTPDKRYGSOPELYLEPGWFGSVRRRTTV 240
DB 181 LRLNTGGLTYVDRDINMTVFLQPDVGTTPDKRYGSOPELYLEPGWFGSVRRRTTV 240
QY 241 NCELMDFARSNPPDFEFTVATGTVMSPFWSGEDDHENKMKHEKPFVSVINNKVVDY 300
DB 241 NCELMDFARSNPPDFEFTVATGTVMSPFWSGEDDHENKMKHEKPFVSVINNKVVDY 300
QY 301 QNRGTVPGLKTRIFLDREEVYLSWEKHLKMSYCPYTLWKAIFYNGIOTHSYGVHVFAND 360
DB 301 QNRGTVPGLKTRIFLDREEVYLSWEKHLKMSYCPYTLWKAIFYNGIOTHSYGVHVFAND 360
QY 361 ITASFTTSKEDMKFNTTYHCLNEIKAEIEKKYAKVNSTHSGYDLKYPKTDGGLVW 420
DB 361 ITASFTTSKEDMKFNTTYHCLNEIKAEIEKKYAKVNSTHSGYDLKYPKTDGGLVW 420
QY 421 QPLIQNLRLDAKKNLNNTYSRRRQAEISTDPMEMTNGAGGEYSSENSITVAQVY 480
DB 421 QPLIQNLRLDAKKNLNNTYSRRRQAEISTDPMEMTNGAGGEYSSENSITVAQVY 480
QY 481 AYDNLRIINNILEDLSKACREOHRALVWNLKINPTSVMSMIYVRPVSAKRIGDVI 540
DB 481 AYDNLRIINNILEDLSKACREOHRALVWNLKINPTSVMSMIYVRPVSAKRIGDVI 540
QY 541 SVSNICVVDDTSVSLHSLRLSASDEKCFSPRPVTPKFWNDSTIYKQGLGVNNEILLTT 600
DB 541 SVSNICVVDDTSVSLHSLRLSASDEKCFSPRPVTPKFWNDSTIYKQGLGVNNEILLTT 600
QY 601 TYLETCENTEYFQAQTDVMIYKNYEHLKTVPLSSITTLDTFTALNFTLLENVDFKVE 660
DB 601 TYLETCENTEYFQAQTDVMIYKNYEHLKTVPLSSITTLDTFTALNFTLLENVDFKVE 660
QY 661 LYTRDEXRLNVDFIETMFREYNYAQRVSGRLKDLDLSTNRNQFVDVAFGSLMDDLGA 720
DB 661 LYTRDEXRLNVDFIETMFREYNYAQRVSGRLKDLDLSTNRNQFVDVAFGSLMDDLGA 720
QY 721 GQTVNAVSGVATLFFSSIVTGFINFKNPFGGMLIIVLGVLPFAIYFLTKTKIYETAP 780
DB 721 GQTVNAVSGVATLFFSSIVTGFINFKNPFGGMLIIVLGVLPFAIYFLTKTKIYETAP 780
QY 781 IKMIYPEIDKLKREGKSEIAPISEEELERIVLAMHIHQONSHMETKTRDKPXDLSILTRA 840
DB 781 IKMIYPEIDKLKREGKSEIAPISEEELERIVLAMHIHQONSHMETKTRDKPXDLSILTRA 840

QY 841 QNMLRKRSYGYSNLKNAESVEMLNTL 865
DB 841 QNMLRKRSYGYSNLKNAESVEMLNTL 865

RESULT 2
AAW26475
ID AAW26475 standard; Protein; 845 AA.

XX AC AAW26475;
XX 02-JAN-1998 (first entry)
XX KSHV glycoprotein B.
XX KSFV; gamma herpes virus; glycoprotein B; vaccine; infection;
XX diagnosis.
XX Human Kaposi's sarcoma-associated herpesvirus.
XX Key Location/Qualifiers
FT Misc-difference 841 /label= Leu, Pro
FT /note= "encoded by CYG"
XX

PN WO9712042-A2.

XX 03-APR-1997.

XX 26-SEP-1996; 96WO-US15702.

XX 26-SEP-1995; 95US-0004297.

XX (UNIW) UNIV WASHINGTON.

XX Bosch ML, Rose TM, Strand K;

XX WPI; 1997-212901/19.

XX N-PSDB; AAT84648.

XX DNA encoding glycoprotein B of retroperitoneal fibromatosis and
XX Kaposi's sarcoma associated herpes viruses - useful in vaccines for
XX treatment of herpes infection or for detection of viral DNA

XX Claim 11; Fig 19; 138pp; English.

XX This polypeptide comprises the full-length glycoprotein B (GB) of
XX human Kaposi's sarcoma-associated herpes virus (KSHV). It is the
XX encoded product of a consensus sequence (see AAT84648) obtained from
XX Kaposi's sarcoma tissue DNA. Herpes viruses GB molecules are
XX associated with fibroproliferative and neoplastic conditions in
XX primates, including humans. Claimed GB polypeptides (AAW26473-76)
XX can be expressed in transformed host cells for use in vaccines, for
XX the treatment of herpes virus infection, to assess herpes virus
XX infection, to design or screen antiviral agents and to raise
XX antibodies. Claimed antigenic peptides (see AAW26477-86) derived
XX from KSHV and/or retroperitoneal fibromatosis-associated herpes virus
XX can be used to raise monoclonal or polyclonal antibodies.

XX Sequence 845 AA;

Query Match 42.6%; Score 1937; DB 18; Length 845;
Best Local Similarity 48.5%; Pred. No. 4.8e-134;
Matches 382; Conservative 135; Mismatches 237; Indels 34; Gaps 10;

QY 83 FPYRVCSAGSGVDVFRFQTDHVCPCDASDMVHSEGIILLIYKONIIPFMFRVRKYRKVVTTS 142

DB 63 YQFRVCSAITGELFFNLNEQTCPTDKYHQBGLLVYKKNIVPHIFKVRRTKATSV 122

QY 143 TVYNGIYSDSITNQHTFYKSIETPWTETKMDTIYQCFSNLRNTGGNLLTYVDRDDINMTV 202

DB 123 TVYRGLTESAITNKYELPRVPVLYEISHMDSTVQCFSSMKVNVNGENTFTDRDDVNTTV 182


```

QY 203 FLQPDVGVTPDKVYKSGQPELYLEPCWFMGYSYRRRTTNCVCELMDFMARSNPDPFVAT 262
DB 183 FLQPEGLTDNIQRVFSQPIVAEPCWPGIYVRVTNCEIVDMARSABYNYFVTS 242
QY 263 GDTVEMSPWSEDDHKNHKKPWF-VSVINNYKVVDYQNRGTVPGLKTRIFLDRBYT 321
DB 243 GDTVEVSPFCYNSSCSTPSSNKGSLVQVWLNHTVVTYSDRGTSTPQNRIFVETGAT 302
QY 322 LSWKHKLKMSYCPITLWKAIFYNGIQTEHSGSYHVFANDITASFTTSKEDMKERNVTHC 381
DB 303 LSWASEKTTAACPALMKTFRPSIQTHEDSFHFVANEITATFAPTPVANFTDTTSC 362
QY 382 LNEEIKAEIEKKYAKVNSHYSYGLKYPKTDGGLYVWQPLIQRLNDAKKNLNET-- 439
DB 363 LSTDINTNLASKAKLASTHVPNGVQYFHTTGLYLVWQPMASAINLTHAQDSCGNPTSS 422
QY 440 -----YSSRRRQAESTTPMMEMTONGAGGEYSSENSITVAQVQVAYNLIR 488
DB 423 PPSASPMTTSASRRKRSASTA-----AAGGGGSTDN-LSYTLQAFAYDKLRDG 471
QY 489 INNILEDLSKAWCROHRAALVWNELSKINPTSVMSIYNRPVSASAKRIGDVISVNCIIV 548
DB 472 INQVLEELSRACROVDRNLWYELSKINPTSVMTAIYGRVSAKFVGDALISVTECINV 531
QY 549 DQTSVSLHKLRLLSASDEKCFSPRPVTPFKFMDNSTIYKGLGVNNEILLTLYLETCOE 608
DB 532 DQSSVNIHKSRA--TNSKDVCYARPLVTFKFLNSSNLFTGQLGARNEILLTNNOVETCKD 589
QY 609 NTEYFQAKTDMYIKVNEHLKTVPLSSITLDTFIALNFTLLENVDPKVIELYTRDEKR 668
DB 590 TCEHYFTRNETVLVKDYAVLRTINTDITSLNFTALNLSFIQNDKAEIYSSAEKR 649
QY 669 L-SNVFDLETWPREYNYAQRVSGRKLDD-LSTNRNQFVDAFOSLMDDLGAVGQTVN 726
DB 650 LASSVFDLETWPREYNYTHRLAGUREDLDNTDMNKEFERVLDSEIVADLGGIGKTVN 709
QY 727 AVSGVATLFPSSIVTGFNFKNPFGLMIIWIGVLAIFYLTKTKIYETAPIKMIYP 786
DB 710 VASSVVTLCGLSVTGFNFKHPLGLMIIWIAIILFMLSRRNTIAQAPVKMIYP 769
QY 787 EIDKLKEREKSEIAPISEEELERVLAMHIHQNSHMETKTRDKPSILTRAQNLK 846
DB 770 DVDR---RAPPSGGAP-TREEIKNILLGWHQLQOEEROKADDLKKSTPVSFVORTANGLRQ 825
QY 847 R-SGYSNL 853
DB 826 RLRYGKPL 833

RESULT 3
AAB53128
ID AAB53128 standard; Protein; 829 AA.
XX
AC AAB53128;
XX
DT 28-FEB-2001 (first entry)
XX
DE Macaca mulatta rhadinovirus 17577 RRV ORF8 protein SEQ ID NO:13.
XX
KW Macaca mulatta rhadinovirus 17577; RRV; rhesus macaque rhadinovirus;
KW genome; Kaposi's sarcoma-associated herpesvirus; KSHV; interleukin 6;
KW IL-6; macrophage inflammatory protein; MIP; diagnosis; vaccine;
KW cytostatic; anti-HIV; gene therapy; infection; Kaposi's sarcoma;
KW lymphoproliferative disorder; B-cell hyperplasia; lymphadenopathy;
KW splenomegaly; hypergammaglobulinemia; autoimmune haemolytic anaemia.
XX
OS Macaca mulatta rhadinovirus 17577.
XX
PN WO200028040-A2.
XX
PD 18-MAY-2000.
XX

```

```

PF 05-NOV-1999; 99WO-US26260.
XX
PR 06-NOV-1998; 98US-0107507.
PR 20-NOV-1998; 98US-0109409.
XX
XX (UYOR-) UNIV OREGON HEALTH SCI.
PI Wong SW, Axthelm MK, Searles RP;
XX
XX WPI; 2000-376552/32.
XX
PT New rhesus rhadino virus for producing non-human primate model useful
PT for testing potential treatments and efficacy of the candidate vaccine
PT for conditions associated with RRV infection .
XX
XX Claim 5; Page 123-124; 141pp; English.
XX
CC The present invention describes a novel rhesus macaque rhadinovirus
CC called macaca mulatta rhadinovirus 17577 (RRV). AAC64754 represents the
CC RRV genome sequence, and AAB53123 to AAB53204 represent the proteins
CC encoded by the genome sequence. The present invention also specifically
CC claims the individual open reading frame (ORF) nucleotide sequences from
CC the genome which encode the individual proteins, but these sequences are
CC not given. A non-human animal infected with RRV can be used for testing
CC the efficacy of drug in the treatment of condition associated with
CC infection with RRV such as Kaposi's sarcoma, lymphoproliferative
CC disorders, B-cell hyperplasia, lymphadenopathy, splenomegaly,
CC hypergammaglobulinemia or autoimmune haemolytic anaemia, by
CC administering the drug to an immuno-compromised non-human primate
CC preferably Rhesus macaque monkey obtained by as a result of infection
CC by Simian Immunodeficiency Virus (SIV). RRV is useful for producing
CC non-human primate model for testing potential treatments for conditions
CC associated with RRV infection. It is also useful for testing the
CC efficacy of the candidate vaccine against RRV infection or conditions
CC associated with its infection by administering the vaccine to the
CC subject capable of infection with RRV, inoculating the subject with RRV
CC and observing the effect of vaccine. AAC64755 to AAC64765 and AAB53205
CC to AAB53213 represent sequence used in the exemplification of the
CC present invention.
XX
SQ Sequence 829 AA;
XX
Query Match 42.4%; Score 1927.5; DB 21; Length 829;
Best Local Similarity 44.4%; Pred. No. 2.3e-133;
Matches 388; Conservative 146; Mismatches 258; Indels 81; Gaps 14;
QY 11 VLALWLYQVALYSLSAETGVTPPTATWSTESPLTGHYTHDSSHGREGNNENRDS 70
DB 10 LLRAWVV--IIAIGTAVGE-NVITPKGATTAKTP----- 42
QY 71 EQKNVIYGSPT-----FPYRVCASGVGVFRFQTDHVCPSDASDMVHSEGIL 118
DB 43 -----GPSTPTPPENPRAEAFKFRVCASASATGELFRNLEKTCFGTEKTHQEGIL 94
QY 119 LIYKQNIIPFMRVRKRVKVVTTSTVYNGIYSDSTNQHTFYKSIPEWTEKMDTIYOCF 178
DB 95 MVFKKNIVPHFKVRVRYRKVATSVTVRGVTAVTGKQEVIRVPVQYENHMDITTYOCF 154
QY 179 NSLRNTGGLLTYVDRDDINMTVFLQPDVGTDPVKRYGSPQPELYLEPGWFGWSYRRRT 238
DB 155 SSMRVNVNGIIVNTYDRTDFTNQTVFLQPEGLTDNIQRYFSQPVLVYTPGWFPGIYVRVT 214
QY 239 TVNCELMDMARSNPDPDFVTATGDTVEMSPWSE-----DDHKNHKKHKKPWFVSVI 292
DB 215 TVNCEIVDMTARSABYNYFVLTALGDTVEVSPFCYNSSCSTPSSNKGSLVQVWLNHTVVTYS 269
QY 293 NNYKVVDYQNRGTVPGLKTRIFLDRBYTSLWSKHLKMSYCPITLWKAIFYNGIQTEHSG 352
DB 270 TNYTIVDFATR--QPTTETRVFADSGEYTVSWKAEDPKSAVCALTWKTFPRAIQTH 327
QY 353 SYHFVANDITASFTTSKEDMKERNVTHYCLNEEIKAEIEKKYAKVNSHYSYGLKYPK 412
DB 328 SYHFVANDVATFTTSPLSQVNTFTGTYPCLNDVDTQKTLNATIKKLSDTHTNGSEQYVET 387

```

QY 413 DGGYLWQPLQIQRLLDAKNNLNNETYSSRRRQAESTTDPMMWMTGNGAGGEYSSENS 472
 Db 388 EGGFLWQPLTPLSLADEWRELNGTT-----PAPPTTTSTANVRVRSGVNEQATDD 440
 QY 473 ITVAQVQAYDNLIRINNILEDLKSKAWCREQRAALVNNELSKINPTSMYIYRNPVS 532
 Db 441 LAAPQLQFAYDKURASINKVLEELSRAWCREQVDRDTVMYELSKINPTSMYIYGRPVS 500
 QY 533 AKRIGDVISVNCIVVDQTSVLHKSRLLSASDEKCFSPRPVTFKPMNDSTYKQGLGV 592
 Db 501 AKFVGDAISVDCVAVDQASVSIHKSRL--TSTPGICYSRPPVTFRLNLTTLFKQGLGP 558
 QY 593 NNEILLTTTLETQENTYTFQAKTDMYIKYKVEHLKTVPLSLITLDFIALNPTLLE 652
 Db 559 RNEILLTDNQVEACKETCEHYFTASNVTYYKDFVFKKINTSEISTLGTFIALNUSFIE 618
 QY 653 NVDFKVIELTRDEKRLS--NVFDIETMFREYNYIAQVSGRLKDL---LDLSTNRNQFVD 708
 Db 619 NIDFRVIELYSRAEKKLSGVSFDIETMFREYNYITQRLAGLREDLONTIDL--NRDLAR 676
 QY 709 AFGLSMDLGAAGQTVVNAVSGVATLFPSSIVTGFINKNPPFGGMLMIIVIGVLAIFY 768
 Db 677 DLSEIVADLGDVGRTVVNAVSSVITLFGSIYSGFINKSPFGGMLMILVIVAVLIVFA 736
 QY 769 LTKKTKIYETAPIKMIYPEIDKLEREGKSEIAPISEEELERIVLAWHIHQSHMETKT 828
 Db 737 LNRNTAIAQAPRMVYDIDKQWPSGGK-----VDOEQIKNLAGHQLQQBERRLDE 791
 QY 829 RKDPKDSILTRAQNMRLKR--SGYSNLKNABSEVE 860
 Db 792 QORSAPSLFRASDGLKRRRGYKPLENEBAQE 824

RESULT 4
 AAU14094
 ID AAU14094 standard; peptide; 856 AA.
 AC AAU14094;
 AC AAU14094;
 AC AAU14094;
 DT 21-NOV-2001 (first entry)
 DE Peptide sequence from Epstein-Barr virus glycoprotein precursor gp115.
 DE Anti-retroviral; DP178-like; DP107-like; antifusogenic; antiviral;
 KW Glycoprotein gp110 precursor; gp115; HIV transmission.
 OS Epstein-Barr virus strain B95-8.
 OS WO200151673-A2.
 PN 19-JUL-2001.
 XX 05-JUL-2000; 2000WO-US35727.
 XX 09-JUL-1999; 99US-0350841.
 XX (TRIM-) TRIMERIS INC.
 PA Jeffs P, Lackey JW, Erickson JB, Lawless MK, Merutka G;
 PI WPI; 2001-442157/47.
 DR Identifying a compound that inhibits the formation of or disrupts a
 PT DP107/DP178 complex, especially compounds with antifusogenic, antiviral
 PT or intracellular modulatory activity, by detecting the formation of a
 PT DP107/DP178 complex -
 XX Disclosure; Fig 32; 259pp; English.
 CC The present invention relates to peptides which exhibit anti-retroviral
 CC activity. The peptides of the invention (AAU12559-AAU14009) comprise
 CC DP178-like' and DP107-like peptides. The DP178 peptide corresponds

CC to amino acids 639-673 of the transmembrane protein gp41 from human
 CC immunodeficiency virus 1 (HIV-1) isolate LAI. The DP107 peptide
 CC corresponds to amino acids 558-595 of gp41 from HIV-1LAI. The invention
 CC also relates to a method of identifying compounds that inhibit the
 CC formation of or disrupts a DP107/DP178 complex. The method comprises
 CC detecting the formation of a DP107/DP178 complex, both in the presence
 CC or absence of a test compound, in a reaction mixture containing DP107
 CC and DP178 peptides. The method is useful for identifying compounds,
 CC including small molecule compounds, which may themselves exhibit
 CC antifusogenic, antiviral or intracellular modulatory activity. The
 CC DP178-like/DP107-like peptides are useful to inhibit human and non-human
 CC retroviral, particularly HIV, transmission to uninfected cells. The
 CC present sequence represents a peptide sequence from Epstein-Barr virus
 CC strain B95-8 glycoprotein gp110 precursor, gp115.
 XX
 SQ Sequence 856 AA;

Query Match 40.5%; Score 1840; DB 22; Length 856;
 Best Local Similarity 45.8%; Pred. No. 7.1e-127;
 Matches 370; Conservative 142; Mismatches 234; Indels 62; Gaps 13;

QY 81 STFPYRVCASGVGVFRFQTDHVCPCDASDMV-HSEGILLIYKQNIIPMFRVRKYRKVV 139
 Db 44 TSFPFRVCELSSHGDLFRSSDIQCPSPGTRENHTEGLLMVFKDNIIPYSFKVRSYTKIV 103
 QY 140 TTSTVYNGIYSDITNQHTFYKSIPEWTEKMDTIYQCFNSLRNTGGNLLTYVDRODDIN 199
 Db 104 TWILIYNGWADSVNHRHEEKFSVDSYETDQMDTIYQCFNAVMTKDGUTRVYVDRODGN 163
 QY 200 MTVFLOQVDEGVTDPVKRYGSGQPELYLEPGWFGSVRRRTTVNCELMDMFARNSPPDFV 259
 Db 164 IIVNLKPTGGLANGVRVYASQTELYDAPGMLIWTYTRITTVNCLITDMWAKNSPDDFFV 223
 QY 260 TATGDTVEMSPFWSGEDDHENKMKHPFVSVINNVYVDYQNGRVPIGKTRIFLDREE 319
 Db 224 TTTGTQTVEMSPFYDGK--NKETFHERADSFHVRTNYKIVDYDNRGTPNQGERRAFLDKGT 281
 QY 320 YTLSEKHLKN--MSYCPGLTLWKAIFYNGIOTESHSGSVHFVANDITASFTTSKEDMKEFNT 378
 Db 282 YTLNWK--LENRTAYCPLQHWQTFDSTIATETGKSIHFVDEGTSSFTVNTTVGIELDPA 339
 QY 379 YHCLNEBEIKAEIKYAKVNSTHSHYKGLK-----YFKTDGGLYLVNQPLIQNR 427
 Db 340 FKCIIEQ-----VNKTHEYEAQVDRTYKQCAITVYFITSGLLLAWLPLTPRS 388
 QY 428 LLDANK--KLNNETYS-----RRSRQAESTTDPMMWMTGNGA 463
 Db 389 LATVKNLTELTTTSPSPSPSPAPSAARGSTPAAVLRRRRRRDAGNATTP---VPPTAP 445
 QY 464 GGEYSSENSITVAQVQAYDNLIRINNILEDLKSKAWCREQRAALVNNELSKINPTSMY 523
 Db 446 GKSLGTLNPNPATVQIQFAYDSLRRQINRMGLDLARAWCLEQKRONNVLRELTKINPTVM 505
 QY 524 SMYINRPVSAKRIGDVISVNCIVVDQTSVLHKSRLLSASDEKCFSPRPVTFKPMNDS 583
 Db 506 SSIYKKAFAAKRLGDVISVQCVPVNOATVTLKSNR--VPGSETMYSRPLVSFSFINDT 564
 QY 584 TTYKQGLGVNNEILLTTTLETQENTYTFQAKTDMYIKYKVEHLKTVPLSLITLDTF 643
 Db 565 KTYEGQLGTDNEIFLTKMTEVCQATSYQYFSGNBEIHVYNDYHFKTELOGIATLQTF 624
 QY 644 IALNFTLLENVDPKVIELYTRDEKRLSNVFDIETMFREYNYIAQVSGRLKDL--LSTN 702
 Db 625 ISLNTSLINIDFASLELYSRDEQRAVNFDELGIFREYNFQANJARGKLDLNDVNSG 684
 QY 703 RNQFVDAFGSLMDLGAAGQTVVNAVSGVATLFPSSIVTGFINKNPPFGGMLMIIVIGV 762
 Db 685 RNQFVDGLGELMDSLGSVCQSITNLVSTVGGFLFSSLVSGFISFFKNPFGGMLILVLVAGV 744
 QY 763 LFAIYPLTKTKIYETAPIKMIYPEIDKLEREGKSE---IAPISEEELERIVLAWHIHQ 819
 Db 745 VILVILSTRTRRQMSQOPVQMLYPGIDELAQOQHASGEGGPGINPISKTELQAIMLA--LHE 802

PT antibodies, useful in vaccines
XX Example; Fig 3; 58pp; English.

CC The SQ in AAN70455 is in the opposite orientation to the prototype
CC orientation of the HCMV genome. The inventors claim antigenic
CC determinants which correspond to a part of the protein coded by the
CC HindIII F fragment of the HCMV genome (between 1378 and 4095 bases
CC from the F/D boundary) and/or the HindIII L fragment (228-2456
CC bases from the L/D boundary). The antigens are useful in vaccines
CC against HCMV and for detecting antibodies specific for HCMV.

XX Sequence 906 AA;

Query Match 24.7%; Score 1121.5; DB 8; Length 906;
Best Local Similarity 30.4%; Pred. No. 8.3e-74;
Matches 282; Conservative 186; Mismatches 350; Indels 111; Gaps 25;

QY 14 LMWLYQVALYSLAETGVTSPPTATWSTESPLTGHGTHDSSHGRGNN-ENRD--- 68
DB 5 IWCL--VVCNLCIVCLGAUVSSSTSHATSS-----THNGSHTSRTTSAQTRSVSQ 55
QY 69 -----SEQNKNIYG-----SPSTFPYRVCSASGVDFRFQTDHVC--- 105
DB 56 HVTSSSAVSHRANETIYNTLLKYGVVGVNTKYPYRVCSMAQGTDLIRFERNICTSMK 115
QY 106 PDASDMVHSEGILLIYKONIIPFMRVRKRYKVTTSTVNGIYSDSITNQHTFYKSIIEP 165
DB 116 PINEDL--DEGIMVYKRNIVAHFTKRVYQKVLTFRSAYIYTYLLGSNTSEYVAPPM 173
QY 166 WETEKMDTIYOCFNSRLNTGNNLLTYVDRODI--NMTVFLQPDVGVTPDVKRYGSOPELY 224
DB 174 WEIHHINKFAQCYSSYRVIYGVTFVAYHRDSYENKTWQLIPDDYSNTHSTRIVTVKQDW 233
QY 225 LEPGFWGYSRRRTTVNCELMDMFARSNPPDFEVTATGDTVEMSPFWSGEDDHENKMH 284
DB 234 HSRGSTW-LYRETCLNCLMTITARRSKYPVHFATSTGVDVYISPFYNGTRNASYFGE 292
QY 285 KPWFEVSNVNNYKV-VYQNRGTVP-LGKTRIFLDREYITLSWE-KHLKNMSYCPILTWKA 341
DB 293 NADKFFIPPNYTVISDFGRPNAAPEHRLVAFLEADSVISWDIQDEKNVT-CQLTFWEA 351
QY 342 FYNGIQTEHSGSYHFVANDITASPTTSKEDMKEPNTYHCLNEIKAEIEKKY-AKVNST 400
DB 352 SERTIRSEADSYHFSAKMTATPLSKQEVNMSDSALDCVRDEAINKLOIENWYNQT 411
QY 401 HSKYGDLYKFTDGLYLVWQPLIQNRLDAKNNNETYS---RRSRQAESTTDPMME 457
DB 412 YEKYGNVSVFETSGLVVFWQGIKQKSLVELERLANRSSLNITHRTTRRSTSDNNTLSS 471
QY 458 MTGNAGAGEYSSENSITVAQVQVAYDMLRIRINNILEDLSKANCREQHRAALVWNELSKI 517
DB 472 M-----ESVHNLYVAQLQFTYDTRLGRIYINRALAQIAEAWCVDQRRTLVEFKELSKI 522
QY 518 NPTSVMSIMYNRPVSARKIGDVTLSVNCIVVDOTSVLSHLKSLLSASDESKCFSPRPVTF 577
DB 523 NPSAILSAIYNKPIAPFMDGVLGLACVINTQTSVKVLKDMN-VKESPGRCISRPVPIF 581
QY 578 KPMNDSTIYKQGLGVNNEILLTTTYLETQENTYEYFQAKTDMYIYKNYEHKLTVP LSSI 637
DB 582 NFANSSVVOYQGLCEDNEILLGNHRTBECQPSLKIFAGNSAYEYVDYLFKRMIDLSI 641
QY 638 TTLDTFTALNFTLLENVDFKVIETYRDEKLSNVDFIETMPREYNYYAORVSGLRKDLL 697
DB 642 STVDSMALDIDPLENTDFRVELYLSQKELSSNVDFLEIEIMREFNSYKORVRYVEDKV 701
QY 698 DLSTNRHQFVDAFGSLMDDLCAVGQTVVNAVSVGATLFFSSIVTGFINFINKPFGGLMII 757
DB 702 D---PLPYLKLGLDMLSGLAGAKAVGAIVAGVGAVSVGEVATFLKNPFGARTIIL 758
QY 758 VVIGLVFAIYFLTKTKIYETAPIKMIYPEI-----DKLKE----- 793
DB 759 VAIAVVIITLIYTRQRLCTQPLQNLFPYLVASDGTVTTSGETKDTSLQAPPSYBESVY 818

QY 794 ---REG-----KSEIAPISEEELERIVLAM-----HIHOONS-----HMETKTR-KD 831
DB 819 NSGRKGGFPSSDASTAAPPTNEQAYQMLLALRLDABQRAQQNGTSLDGTGTQDKG 878
QY 832 PKDSILTRAQNLMLRKRSYGYNLKNASVE 860
DB 879 QKENLLDLRLRH---RKNGYRHLKDSDEEE 904

RESULT 9
AAP91047

ID AAP91047 standard; protein; 906 AA.

XX AAP91047;

XX 04-JAN-1990 (first entry)

XX gB envelope protein of cytomegalovirus.

XX Cytomegalovirus; glycoprotein gp. 55; glycoprotein gB;
XX vaccine; prophylactic agent.

XX Cytomegalovirus.

XX WO8907143-A.

XX 10-AUG-1989.

XX 26-JAN-1989; 89WO-US00323.

XX 29-JAN-1988; 88US-0149715.

XX (CHIR) CHIRON CORPORATION.

XX Spaete RR, Pacht CA;

XX WPI; 1989-249012/34.

XX New recombinant polypeptide from gB gene of cytomegalovirus - useful in
XX vaccines, also diagnostic DNA coding sequences and derived antibodies.

XX Claim No. 1; fig No. 2; 64pp; English.

XX The gB envelope protein derived from glycoprotein gp. 55 contains an
XX epitope immunologically identical to one encoded by the CMV genome, and
XX is useful in vaccines and as a prophylactic agent against CMV
XX infections, or for detecting CMV and CMV Abs. See also AAN90528.

XX Sequence 906 AA;

Query Match 24.5%; Score 1114; DB 10; Length 906;

Best Local Similarity 30.7%; Pred. No. 3e-73;

Matches 284; Conservative 184; Mismatches 355; Indels 102; Gaps 25;

QY 14 LMWLYQVALYSLAETG---VTSPTATWSTESPLTGHGTHDSS--HGERGNENRDS 69

DB 5 IWCL--VVCNLCIVCLGAUVSSSTTRGTSATHS---HHSHHTTSAHSRSGSVQRVT 58

QY 70 BEQ-----NKNIYG-----SPSTFPYRVCSASGVDFRFQTDHVC---PDA 108

DB 59 SSQTQVSHGVNETIYNTLLKYGVVGVNTKYPYRVCSMAQGTDLIRFERNICTSMKPIN 118

QY 109 SMDVHSEGILLIYKONIIPFMRVRKRYKVTTSTVNGIYSDSITNQHTFYKSIIEPWET 168

DB 119 EDL--DEGIMVYKRNIVAHFTKRVYQKVLTFRSAYIYHTYLLGSNTSEYVAPPMWEI 176

QY 169 EKMDIYOCFNSRLNTGNNLLTYVDRODI--NMTVFLQPDVGVTPDVKRYGSOPELYLEP 227

DB 177 HHINSHSQCYSSYRVIYGVTFVAYHRDSYENKTWQLMPDDYSNTHSTRIVTVKQWHBR 236

QY 228 GWFWSYRRRTTVNCELMDMFARSNPPDFEVTATGDTVEMSPFWSGEDDHENKMKHEKPW 287

```

Db 237 RSTW-LYRETCLNLCMVTTITARSKYPHFFATSTGDDVVDISPFYNGTNRNRSYFGENAD 295
Qy 288 FVSINNYKVVDYQNRGTVPGLKTRI--FLDREETYLSWE-KHLKNMSYCPLTLWKAFYN 344
Db 296 KFFIFPNYTVISDFGRPNSEALETHRLVAFLEADSVISWDIOEKVNT-CQLTFWEASER 354
Qy 345 GLOTHSGSYHVFANDITASFTTSKEDMKFNTTYHCLNEEIKABIEKKY-AKVNSTHSHK 403
Db 355 TIRSEADSYHFSACKMTATFLSKQEVNMSDALDCVRDEAINKLOQIFNTSYNQTYEK 414
Qy 404 YGDLXVKTDDGLYLWVOPLIQNRLLDAKLNNETYSRSPROASTITDPMENTGGA 463
Db 415 YGNVSFETGTGLVFWQIQKOSLVELERLANRSLNLTNRKST-----DGNNA 467
Qy 464 G--GEYSSENSITVAQVQVAYONLRIRINNILEDLSKAWCREQRAALVWNELSKINPTS 521
Db 468 THLSNMESVHNLVYAQLOFTYDTRGYNRLAQAIAEAWCVDRRTLEVFELSKINPSA 527
Qy 522 VMSMIYNRVPSAKRGIDVISVNCIVVDQTSVSLHKSRLLSASDEKCFSPRPVTFKFMN 581
Db 528 ILSAIYNKPIAARFMDVLGLASCVTINQTSVKLRDMN-VKESPGRCYSRPPWIFNEAN 586
Qy 582 DSTIYKGLGVNNEILLTTTLYETCOENTYEFQAKTDMYIYKNYEHKTVPLSSITILD 641
Db 587 SSYVQVGLGEDNEILLGNHRTBECQLPSLKIFAGNSAYEYVDYLFKRMIDLSSISTVD 646
Qy 642 TFIALNFTLLENVDFKVIELYTRDEKRLSNVFDIETMFEYNYAQRVSGLRKDLDLST 701
Db 647 SMIALDIDPLENTDFRVLALYSQKELRSNVFDEEIMREFNSYKQKVYVEDKVD--- 703
Qy 702 NRNQVDFAPGMDLDLGVQTVVNAVSGVATLFSIVTGFINFKNPFGGMLMIIVVIG 761
Db 704 PLPPYLKGLDLDLMSGLGAAGKAVGAIGAVGAVSGVATFLKNPFGAFTIILVAIA 763
Qy 762 VLFAYFLTKTKIYETAPIKMIYPEI-----DKLKE-----R 794
Db 764 VIIILYMTQRRLCMOPLQNLFFYLVASDGTVTGNTKDTSLQAPPSYEESVYNSGR 823
Qy 795 EG-----KSEIAPISEELEERIVLAM-----HIHQONS---HMETKTR-KDPKDS 835
Db 824 KGPSPSSDASTAAPPYTNEQAYQMLLALVRLDAEQRAQNGTDSLQGTGTQDKGQKPN 883
Qy 836 ILTRAQNMLRKSGSYNLSKNAESVE 860
Db 884 LLDRLRH---RKNGYRHLKDSDEE 905

RESULT 10
AAB31678
ID AAB31678 standard; Protein; 195 AA.
AC AAB31678;
XX
XX
XX
DT 30-APR-2001 (first entry)
XX
DE A viral glycoprotein B envelope protein fragment.
XX
KW Glycoprotein B; envelope protein; vaccine; passive immunity; pig.
XX
OS Swine gamma herpesvirus.
XX
PN WO200103733-A1.
XX
XX 18-JAN-2001.
XX
XX 07-JUL-2000; 2000WO-US18647.
XX
XX 08-JUL-1999; 99US-0142736.
XX
XX 02-DEC-1999; 99US-0168532.
XX
XX (BIOT-) BIOTRANSPLANT INC.
XX
XX Patience C;

```

```

XX WPI; 2001-103033/11.
DR N-PSDB; AAF25178.
XX
XX New polynucleotide encoding a polypeptide corresponding to porcine
PT gamma-herpesvirus glycoprotein B (pGHV-gpB) for immunizing a pig
PT against a porcine gamma-herpesvirus -
XX
XX Disclosure; Page 67; 82pp; English.
XX
XX The present sequence represents a fragment of a swine gamma herpesvirus
CC glycoprotein B envelope protein. Probes and primers derived from DNA
CC encoding the glycoprotein B envelope protein are used for detecting the
CC presence of swine gamma herpesvirus. Antibodies directed against the
CC envelope protein are also useful for detecting presence of the virus.
CC The envelope protein can be used as a vaccine for creating passive
CC immunity in a pig.
XX
XX Sequence 195 AA;
XX
XX Query Match 22.0%; Score 1002; DB 22; Length 195;
XX Best Local Similarity 100.0%; Pred. No. 5.8e-66;
XX Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 484 NLRIRINNILEDLSKAWCREQRAALVWNELSKINPTSVMSMIYNRVPSAKRGIDVISVS 543
Db 1 NLRIRINNILEDLSKAWCREQRAALVWNELSKINPTSVMSMIYNRVPSAKRGIDVISVS 60
Qy 544 NCIVVDQTSVSLHKSRLLSASDEKCFSPRPVTFKFMNDSTIYKGLGVNNEILLTTTLYL 603
Db 61 NCIVVDQTSVSLHKSRLLSASDEKCFSPRPVTFKFMNDSTIYKGLGVNNEILLTTTLYL 120
Qy 604 ETCOENTYEFQAKTDMYIYKNYEHKTVPLSSITILDFTALNFTLLENVDFKVIELYT 663
Db 121 ETCOENTYEFQAKTDMYIYKNYEHKTVPLSSITILDFTALNFTLLENVDFKVIELYT 180
Qy 664 RDEKRLSNVFDIETM 678
Db 181 RDEKRLSNVFDIETM 195

RESULT 11
AAR77024
ID AAR77024 standard; Protein; 879 AA.
XX
XX AAR77024;
AC AAR77024;
XX
XX 20-APR-1996 (first entry)
XX
XX Canine herpesvirus gB homologue.
XX
KW CHV; glycoprotein gB; vector; attenuation; poxvirus; vaccinia virus;
KW canarypox virus; ALVAC; vCP320; dog; puppy; immunisation; antigen;
KW vaccine.
XX
OS Canine herpesvirus.
XX
XX Key Location/Qualifiers
XX Region 725..741
XX Region /label= Transmembrane_region
XX Region 747..771
XX Region /label= Transmembrane_region
XX
XX WO9526751-A1.
XX
XX 12-OCT-1995.
XX
XX 30-MAR-1995; 95WO-US03982.
XX
XX 29-MAR-1995; 95US-0413118.
XX
XX 30-MAR-1994; 94US-0220151.
XX
XX (VIRO-) VIROGENETICS CORP.

```


Db 744 LHELREYDI--DRVVKTDGNMIMRGLANFPQGLGAVGAQVTVLGAAGAALSTVSGIA 801
 QY 744 NFIKNPFGMLMIIVVIGVLFYFLTKTKIVETAPIKMIYPIDKLKEREGKSEIAP 803
 Db 802 SFIANPFGALATCLLVGLVAFLAYRYSIRLSNPMKALPITTRALKDDARGATAPG 861
 QY 804 SEEE-----LER---IVLAMIHQONSHMETKTRKPKDS--ILTR-AQNMLRKRS--G 849
 Db 862 EEEEFDAKLEQAREMIKYMISLVSVERQEHKAKSKNGGKGLLATRLTLQALRRAPPE 921
 QY 850 YSNLKNAE 857
 Db 922 YQOLPMAD 929

RESULT 15

AAR14665
 ID AAR14665 standard; Protein; 904 AA.

XX AAR14665;

XX AC

XX 25-MAR-2003 (updated)

DT 23-JAN-1992 (first entry)

XX HSVGB polypeptide.

DE Vaccine; antigen.

XX Herpes simplex virus.

OS JP03218397-A.

XX 25-SEP-1991.

XX 21-JUN-1990; 90JP-0161448.

XX 30-NOV-1989; 89JP-0308941.

PR 22-JUN-1989; 89JP-0158238.

PR 21-JUN-1990; 90JP-0161448.

XX (TAKE) TAKEDA CHEM IND LTD.

XX WPI; 1991-328397/45.

DR N-PSDB; AAQ14478.

XX HSVGB polypeptide(s) obtd. by recombinant DNA techniques -

PT useful as vaccines against HSV and in diagnosis, can be produced

PT cheaply and safely.

XX Claim 1; Fig 2; 24pp; Japanese.

XX The sequence was deduced from DNA and is that of the HSVGB polypep-

CC tide. The recombinant protein or esp. truncated forms of it, can be

CC used to prepare vaccines for prophylaxis of HSV infection and for use

CC in diagnostic kits.

CC See also AAR14666.

CC (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 904 AA;

SQ

Query Match 18.6%; Score 847.5; DB 12; Length 904;

Best Local Similarity 27.7%; Pred. No. 1.4e-53;

Matches 244; Conservative 180; Mismatches 362; Indels 95; Gaps 27;

QY 28 ATGVTSPNTATWSTESP--LTGHYGTDSHGE-RGNENRDEEQNKNTYGSPTFP 84

Db 62 APTGDKPKKKNKPNPPPPRAGDNATVAAGHATLREHLRDIKAKNTDANFY----- 114

QY 85 YVCSASGVDFPRQTDHVCPSDQD-MVHSGLILLIYKQNIIPFMRVRKYKVVTTST 143

Db 115 --VCPPTGATVQFEQPRRCPTREGQNYTEGIAWFKENIAPYKFKATMYKDVTSQ 172

QY 144 V-----YNGIYSDSITNQHTFYKSIPEWE--TEKMDTIYQCFNSURLNTGNNLLTYV 193
 Db 173 VMFGHRYSQFNGIIFED-----RAPVPEEVIDKINAKGVCRSTAKYVRNNLETTAF 223
 QY 194 DRDDINMTVFLOPDGV-----TPDVKEYGSOPELYLEPGHFWGYSRRRTTVNCBLM 245
 Db 224 HRDDETMDLKPANAATRTSRGWHTTDLKYNPSRVEAF-----HRYGTTVNCIVE 274
 QY 246 DMFARSNPPFFVFTATGDTVEMSPFWS-GEDDH-ENKMHKPFVSVINNY-KVVDYQN 302
 Db 275 EVDARSVPYNEFVLATGDFVMSPPFYCYREGSTEHTSYAADRFKQVDGYARDLITKA 334
 QY 303 RGTPLGKTRIFLDREYYTLSWEKHLKNMXYCPLTKMAFYNGIQTEHSGSYHFVANDIT 362
 Db 335 RATAP--TTRNLLTTPKFTVAMDVPRKPSVCTWTKQOEVDMLRSEYGGSFSSDAIS 392
 QY 363 ASFTT-----SKEDMKFEFTVTHCLNEETKAEIKKYA-KVNSTHSKYGDLKYFKTDG 414
 Db 393 TTFNTLTYPLSRVDLGD-----CIGKDARDAMDRIFARRYNATHIKVGQPOYYLANG 446
 QY 415 GLYLWQPLIQNRLLDKKNLNNETYSRRSROAESTTDPMMEMTNGCAGGEYSSENSIT 474
 Db 447 GLIAYQPLLSNTLAE-----LYVREHLREQRKPNPTPPPGASANASVERIKTSSIE 502
 QY 475 VAQVQAYDNLRIIRINILEDLSKAWCREQHRALVWNLKINPTSYMSIYNRPVSAR 534
 Db 503 FARLQFTYNNHQRHVNDMLGRVAIAWCELQNHLELTWNEARKLNPATASVTVGRRVSAR 562
 QY 535 RIGDVISVNCIVVDQTSVSLHLSLRLSASDEKCFRPPVTFKPMNDSTIYKQOLGVNN 594
 Db 563 MLGDVMAVSTCVPAADNVIQNSMR-ISSRPGACYSRPLVSFRYEDQGLVEGOLGNN 621
 QY 595 EILLTTTTLTQENTTEYFOAKTDMYIKNYEHLKTVPLSSITTLDTFIALNPTLLENV 654
 Db 622 ELRUTRAIEPCTVGHRRYFTGGYVYFEYASHQLSRADITTVSTFDLNTIMLEDH 681
 QY 655 DFKVIELYTRDEKRLSNVFDIETMFREYNYVAQVSLRGLKDLDLSTNRNOF--VDAFGS 712
 Db 682 EFVPLEVYTRHEIKDGLDYTEVORRNQLHDLRFADI-DTVIHADANAAMFAGLGAFFE 740
 QY 713 LMDDLGL-AVGQTVNAVSGVATLFSIVTGFINIKPFGGMLMIIVVIGVLFALYFLTK 771
 Db 741 GMGLGRAGVKVMGIYGGVV-----SAVSGVSSFSMNPFGALAVGLLVLAGLAAFAFR 796
 QY 772 KTKIVETAPIKMIYPIEDK-LK-----EREKGSIAPISEB-----ELERIVLAMHIHQON 821
 Db 797 YVMLQSNPMKALYPLTTKELKNTNPDASGEGBEGGDFDEAKLAERIMRYMALVSAM 856
 QY 822 SHMETKTRKDPKDSILT-RAQNML---RKRSGYNLKNKNAES 858
 Db 857 ETEHKAKKGKTSALLSAKVTDVMWRKRNTNYTQVPNKDS 897

RESULT 16

AAP71135

ID AAP71135 standard; Protein; 903 AA.

XX AAP71135;

XX 25-MAR-2003 (updated)

DT 30-APR-1991 (first entry)

XX Herpes Simplex Virus-1 GB from pKEXX.

XX Vaccine; prophylaxis; treatment; Herpes Simplex Virus-1;

XX glycoprotein; GB.

XX Herpes simplex virus type 1 (KOS).

XX US4642333-A.

XX 10-FEB-1987.

Query Match 18.6%; Score 844; DB 23; Length 904;
Best Local Similarity 28.4%; Pred. No. 2.5e-53;
Matches 247; Conservative 164; Mismatches 366; Indels 92; Gaps 27;

Qy	33	TSPPNTATWSTESPLTGHYTHDSHGREGNNENRDSEQNKNIIYGSPTSPFYRVCSAGS	92
Db	76	TPPPD-----ANATAAGSHATLR--AHLREIKVENADAQ-----FYVCPPT	115
Qy	93	VGDVFRFQDTHVCPDASD-MVHSEGIILLIYKQNIIPFMRVRYKRVVTTSTV-----	144
Db	116	GATVVQFQPPRCPTRPQGNVTEGIAVFKENIAPYKFKATMYKDVTVSQVWFGRHS	175
Qy	145	-YNGIYSDSITNQHTFYKSIPEWE--TEKMDTIYOCFNSLRNLNTGNLLTYVDRODINMT	201
Db	176	QFMGIFD-----RAPVPFEEVIDKINTKGVCSTAKYVRNMWETTAHRDDHETD	226
Qy	202	VFLQPVGDVTPDVRYKYSQPELYLEPGFWGSYRRRTTNCBLMDMFARSNPFPFFVTA	261
Db	227	MELKPAK-VATRTSRGWHTTDLKYNPSRVEAHRVYGTIVNCIVVEEDARSVPYDEFLA	285
Qy	262	TGDTVEMSPFWS-GEDDH-ENKMHKPFVSVINNY-KVVDYQNRGTVPGLKTRFLDRE	318
Db	286	TGDFVYMGPFPGYREGSHTETSAAADRFPQVDGFYARDLTTTKARATSP--TTRNLLTTP	343
Qy	319	EYTLSEWHLKNMSPCLTWKAFYNGIQTEHSGSYHFVANDITASFTT-----SKED	371
Db	344	KFTVANDVVKPACTWTKQWEDEMLRAEYGGSRFSSDAISTFTTNLTTEYSLSRVD	403
Qy	372	MKEFNTTYHCLNEETKABIEKKYA-KVANSTHSGYGLKYFKTDGGLYLWQPLIQNRLD	430
Db	404	LGD-----CIGRDAREAI DRMFARKYNATHIKVGOPOVYLATGGFLIAYQLLSNTLAE	457
Qy	431	AKNKLNNETYSR-RSRQAEASTTDPMMEM-TONGAGGEYSSENSITVAQVQVAYONLRIR	488
Db	458	---LYVREYMRQDRKPRNATPAPLRAPSANASVERIKTTSSIEFARLQFTYNIQRH	513
Qy	489	INNILEDLSKAWCREOHRAAVLWNELSKINPTSVMSMIYNRPVSAKRGDVTLSNCIVV	548
Db	514	VNDMLGRITAVWCELQNHLETLWNEARLKNPAIASATVGRRVSRMLGDVMAVSTCPVP	573
Qy	549	DOTSVLSHKSRLLSASDEKCFSRPPVTFKFMNDSTIYKQGLGVNNEILLTTTYLETQOE	608
Db	574	APDNVIVQNSMR-VSSRPGTCYSRPLVSFRYEDQGLLEGQLGENNELRLTRDALEPCTV	632
Qy	609	NTEYFFQAKTDMYIYKNEYHLKTVPLSSITTLDTPIALNFTLLENVDFKVIELYTRDEKR	668
Db	633	GHRRYFIFGGGVVPEEYAYSHQLSRADVTTVSTFDLNIITMLEDHEFVPLEVYTRHEIK	692
Qy	669	LSNVFDIETMPREYNYYAQRVSGRLKDLLDLSTNRNQVD--AFGSLMDDLG-ANGQTVV	725
Db	693	DSGLDYTEVORRNQHLDRFADI-DTIVRADANAAMFAGLCAPFEGMGDLGRAVGKVMV	751
Qy	726	NAVSGVATLFSISITGTFINIKNPPFGMLMIITWIGLVFAIYFLTKTKIYETAPIKMIY	785
Db	752	GVVGGV---SAYSGVSSFNPNFGALAVGLLVLAGLVAFAFFRYVLQLOLRNPKALY	807
Qy	786	PEIKLKEREGKSEIAPISBE-----ELERIVLAMHTIQQNSHMETKTRKDP	832
Db	808	PLTTKELKLTSDPGVGGECECAEGGFEAKLAEBAREMIRYMALVSAAMERTEHKARKKG	867
Qy	833	KDSLIL-TRAQNM-LRKR--SGVSNLKNAE	857
Db	868	TSALLSSKVTNMLVRKNRKARYSPLHNE	896
RESULT	18		
AA	AA	AA	AA
ID	AA	AA	AA
XX	XX	XX	XX
XX	XX	XX	XX
XX	XX	XX	XX
DT	25-MAR-2003	(updated)	
DT	10-MAR-2003	(updated)	
DT	30-APR-1991	(first entry)	
XX	Recombinant herpes simplex virus.		
XX			

KW	Herpes simplex virus; E.coli; plasmid; vector; promoter; antigen;
KX	vaccine; glycoprotein; HSV gB.
XX	
OS	Herpes simplex virus.
XX	
PN	JP62115288-A.
XX	
PD	26-MAY-1987.
XX	
PF	14-NOV-1985; 85JP-0255971.
XX	
PR	14-NOV-1985; 85JP-0255971.
XX	
PA	(KAGA) KAGAKU OYOBI KESSEI RYOHO.
XX	
DR	WPI: 1987-183377/26.
XX	
DR	N-FSDB; AAN70681.
XX	
PT	Recombinant DNA contg. herpes simplex virus gene - is integrated in
PT	E.coli plasmid vector, contg. incorporated promoter region for
PT	animal cells, giving antigen.
XX	
PS	Disclosure; Fig 3(1-3); 13pp; Japanese.
XX	
CC	The gene encoding HSV gB is cloned into E.coli. To increase the
CC	efficiency of expression, the gene encoding the C-terminal region
CC	of gB, which is very hydrophobic, is removed and the remaining gene
CC	is cloned into an expression vector. The vector is composed of pBR322,
CC	SV40 early promoter and the HSV gB gene.
CC	The construct is useful for the production of a good antigen for
CC	a HSV vaccine.
CC	(Updated on 10-MAR-2003 to add missing OS field.)
CC	(Updated on 25-MAR-2003 to correct PA field.)
XX	
QQ	Sequence 903 AA;
XX	
Query Match	18.6%; Score 843.5; DB 8; Length 903;
Best Local Similarity	27.6%; Pred. No. 2.7e-53;
Matches 245; Conservative 180; Mismatches 361; Indels 101; Gaps 28;	
Qy	24 SLSTAETGVTSPPNTATWSTESP--LTGHYGTDDSSHGE-RGNNENRDSBEQNKNIYGSP 80
Db	57 ALGAAPTGDPKPKNKPKPKNTPPRPAGDNATVAAGHATLREHLRDIKAENTDANFY--- 113
Qy	81 STFPYRCSASGVGDVPRFQTDHVCDPASP-MVHSEGILLIYKONIIPMFVRVKYKVV 139
Db	114 -----VCPPTGATVVQFQPRKCRTPRPGQNTVEGIAVVFKENIAPYKPKATMYKDV 167
Qy	140 TTSTV-----YNGIYSDSITNQHTFYKSIPEWE--TEKMDTIYOCFNSLRINTGGNL 189
Db	168 TVSQVWFGRHYSQPMGIFED-----RAPVPFEEVIDKINAKGVCSTAKYVRNNLE 218
Qy	190 LTVYVDRDDINMTVLPQVDGV-----TPDVKRYGSGQPELYLEDPGFWGVSYYRRRTTVN 241
Db	219 TTAAPHRDDHETMELKPNAAATRTSRGWHHTDLKYNPSRVEAF-----HRYGTTVN 269
Qy	242 CELMDMEFARNPPDPFVTATGDTVENSPWS-GEDDH-ENKMHKEKPFVSVINNY-KVV 298
Db	270 CIVEEDARSVPYDPEFVLATGDFVYNSPFYGRSGHTEHTTVAADRFKQVDGFYARDL 329
Qy	299 DYQNRGTVPLGKTRIFLDREBYTLSEKHLKNMSYCPDLTWKAFYNGIQTEHSGSYHFVA 358
Db	330 TTKARATAP--TIRNLLTTPKFTVANDWVPKRPSVCTMTKQWDEMLRSEYGCSPFSS 387
Qy	359 NDIATSTT-----SKEDMKFENTTYHCLNBEIKAEIEKKYA-KVNSTHSKYGDUKYF 410
Db	388 DAISTTTTNNLTVEPLSRVDLGD-----CIGKDARDAMDRIFARRYNATHIKVGQPOYY 441
Qy	411 KTDGGLYLWQPLIONLLDAKNNLNNETYSRRSRQAESTTDPMMETNGAGGEYSSE 470
Db	442 LANGGFLIAYQPLLSNTLAE-----LYVREHLREQSRKPNPTPPPGASANASVERIKTT 497
Qy	471 NSITVAOVAYDNLRIIRINILEDLSKWCREOHRALVNNELSKINPTSVMSMINRP 530

Db 498 SSIEFARLQPTYNHQRHNDMLGRVAIAWCELQNHLETLWNEARKLNPNAIASATVGR 557
Qy 531 VSAKRIGDVSVCNIVVDQTSVSLHSLKSLRLSASDEKCFSPRPVTFKFNDSITTYKGL 590
Db 558 VSARMLGDMVAVSTCVPAADNVVQNSMR-ISSRPGACYSRPLVSFRYEDQGPLVEGOL 616
Qy 591 GVNNEILLTLYETCQENTYEFQAKTDMVIYKNYEHKTVPLSSITTLDTFIALNFTL 650
Db 617 GENNELRLTRAIEPCTVGHRRYFTFGGVYFPEYAYSHQLSRADITTVSTFIDLNTM 676
Qy 651 LENVDKVIELYTRDEKRLSNVFDIETMFREYNYYAQRVSGLRKDLDSLSTNRNQF--VD 708
Db 677 LEDHEFPVLEVYTRHEIKDGLDYTEVQRNQLHDLRFADI-DTVIHADANAAMPAGLG 735
Qy 709 AFGSLMDLIG-AVCQTVNAVSGVATLFSISIVTGFINFKNPFGGMLIIVIGVLPAY 767
Db 736 AFFEGMDLGRAVGKVMVIGGVV----SAVSGVSMFNPFGALAVGLLVLAGLAAAF 791
Qy 768 FLTKTKIYETAPIKMIYPEIDK-LK-----EREGKSEIAPISE-EELERIVLA 814
Db 792 FAFYVMRLQSNPKALYPLTKELKNPTPDASGEGBEGDFDEAKLAAREMIRYMAL 851
Qy 815 MHIHQNSHMETKTRKDPKDSILT-RAQNML---RKRSGYSNLKNAE 857
Db 852 VSAMERTEH---KAKKGTSLRLSAKVTDVMWRKRRNTNYQVFNKD 895

RESULT 19
AAW00375 ID AAW00375 standard; Protein; 904 AA.
XX AC AAW00375;
XX DT 24-JAN-1997 (first entry)
XX DE HSV-1 glycoprotein B.
XX KW HSV-1; glycoprotein B; vaccine; diagnosis; immunoassay.
XX OS Herpes simplex virus type 1.
XX FH Key Location/Qualifiers
FT Peptide 1..22
FT Peptide /label= Sig_peptide
FT Peptide 14..119
FT Domain /label= Dominant_type-specific_epitope
FT Peptide 23..744
FT Peptide /label= Extracellular_domain
FT Peptide 295..507
FT Domain /label= Type-specific_epitope
FT Domain 745..798
FT Domain /label= Transmembrane_domain
FT Domain 799..804
FT Peptide /label= Cytoplasmic_domain
FT Peptide 814..901
FT Peptide /label= Cross-reactive_antigenic_epitope
XX W09632962-A1.
XX PD 24-OCT-1996.
XX PF 19-APR-1996; 96WO-US05316.
XX PR 21-APR-1995; 95US-0426604.
XX PA (UYNE-) UNIV NEW MEXICO STATE.
XX PI Bell R, Goade DE, Jenison S;
XX WFI; 1996-485557/48.
XX PT New type-specific and cross-reactive Herpes Simplex Virus epitope (s)

PT - used for diagnosing HSV-1 and/or HSV-2 infections and in vaccines
PT for producing neutralising antibodies
XX Disclosure; Fig 4; 37pp; English.
CC Herpes simplex virus type 1 (HSV-1) glycoprotein B (gB1) (AAW00375)
CC and HSV-2 gB2 (AAW00376) include type-specific and cross-reactive
CC antigenic epitopes. Recombinant type-specific epitopes can be used
CC in immunoassays to distinguish between HSV-1 and HSV-2 infections.
CC Recombinant cross-reactive epitopes are useful for the diagnosis of
CC both HSV-1 and HSV-2 infections. Both types of epitope can be used
CC in vaccines.
XX SQ Sequence 904 AA;
SQ Query Match 18.5%; Score 841; DB 17; Length 904;
Best Local Similarity 27.7%; Pred. No. 4.2e-53;
Matches 242; Conservative 179; Mismatches 345; Indels 108; Gaps 28;
Qy 35 PNTATMSTESPLTGHYTHDSSHGE-RGNENRDSSEQNKNIYGSFSTFPYVCSASGV 93
Db 80 PPRPA-----GNATVAAGHATLRBHLRDIKAENTDANFY-----VCPPTG 121
Qy 94 GDVFRFQTDHVCDDASD-MVHSEGIILYKQNIIPFMFRVRYKRVKVTSTV----- 144
Db 122 ATVVQFEQPRCPTRPEGQNYTEGIAVFKENIAPYKFKATWYKDVTVSQVMFGRHSQ 181
Qy 145 YNGIYSDSTNQHTFYKSIETPWE--TEKMDTIIQCFNSRLNTGGNLLTYVDRDDNMTV 202
Db 182 FMGIFED-----RAPVPFEVIDKINAKGVCSTAKYVRNNLETTAFHRDDHETDM 232
Qy 203 FLOPVDGV-----TPDKRYGSOPELYLEPGWFGWSYRRRTTVCNLCMDMFAFSNPP 254
Db 233 ELKPANAATRTSRGWHITDLKYNPSVEAP-----HRYGTTVNCIVEEDARSVP 283
Qy 255 FDFVFTATGDTVMSPFWS-GEDDH--ENKMKHEKPFVSVINNY-KVVDYQNRGTVPGLKT 311
Db 284 YDEFVLATGDFVYMSPFYVREGSHTHTSYAADRFKQVDGFYARDLTTKARATAP--TT 341
Qy 312 RIFLDREEYTLSEKHLKMSYCPLTLKAFYNGIOTEGSHGSYHFVANDITASFTT----- 367
Db 342 RNLLTTPKFTAWDMVYKRPVSVCTMTKQEVDEMLSEYGGSPFRSSDAISTTFTTNLTE 401
Qy 368 ----SKEDMKEFNTYHCLNEETKAEIEKKYA-KVNSTHSKYGDKLYFKTDGGLYVWOL 423
Db 402 YPLSRVDLGD-----CIQKARDAMDRIAFARYNATHIKVGPQPYLLANGGLIAYQPL 455
Qy 424 IQNRLLDKAKNLNNETYSRRSRQAESTTDPMMWMTGNGAGGEYSSENSITVAQVQYAYD 483
Db 456 LSNLTAE----LVVREHLEQSRKPPNTPPPPGASANASVERIKTTSIEFARLQFTYN 511
Qy 484 NLRIRINNILEDLSKAWCREQRAALVWNLKINPTSVMSIYNPVPAKRGIDVISVS 543
Db 512 HIQRHVNMDLGRVAIAWCELQNHLETLWNEARKLNPNAIASATVGRVRSARMIGDMVAS 571
Qy 544 NCIVVDQTSVSLHSLKSLRLSASDEKCFSPRPVTFKFNDSITTYKGLGVNNEILLTLYL 603
Db 572 TCVPVAAADNVIVQNSMR-ISSRPGACYSRPLVSFRYEDQGPLVEGGLNNEURLRDAI 630
Qy 604 ETCQENTYEFQAKTDMVIYKNYEHKTVPLSSITTLDTFIALNFTLENVDFKVIELYT 663
Db 631 EPCTVGHRRYFTFGGVYFPEYAYSHQLSRADITTVSTFIDLNTMLEDFEVPLEVYT 690
Qy 664 RDEKRLSNVFDIETMFREYNYYAQRVSGLRKDLDSLSTNRNQF--VDAFGSLMDLIG-AV 720
Db 691 RHEIKDGLDYTEVQRNQLHDLRFADI-DTVIHADANAAMPAGLGAFFEGMDLGRAV 749
Qy 721 GQTVNAVSGVATLFSISIVTGFINFKNPFGGMLIIVIGVLPAYLTKTKIYETAP 780
Db 750 GKVMGIVGVV----SAVSGVSMFNPFGALAVGLLVLAGLAAFAFFRYVMRQSNP 805
Qy 781 IKMIYPEIDK-LK-----EREGKSEIAPISE-EELERIVLAMHIHQNSHMETK 827

Db 806 MKALYPLTTKELKUNTPNDPDSAGEGEGDFDEAKLAAREMIRYMALVSAMERTEH---K 862
Qy 828 TRKDPKDSILT-RAQNML---RRKSGYSNLKNAE 857
Db 863 AKKKGTSALLSAKAVTDMVRKERNNTNYTQVPNKD 896

RESULT 20
ID AAB74441 standard; protein; 904 AA.
AC AAB74441;
XX
DT 29-MAY-2001 (first entry)
XX
DE Herpes simplex virus 1 glycoprotein B.
XX
KW HSV-1; HSV-2; glycoprotein B; gB; transmembrane envelope glycoprotein;
KW antigenic epitope; diagnosis; vaccine.
XX
OS Herpes simplex virus type 1.
XX
PN US6197497-B1.
XX
PD 06-MAR-2001.
XX
PF 19-APR-1996; 96US-0632537.
XX
PR 21-APR-1995; 95US-0426604.
XX
PA (UYNE-) UNIV NEW MEXICO STATE.
XX
PI Goade DE, Bell R, Jenison S;
XX
DR WPI; 2001-256360/26.
XX

Continuous, isolated, antigenic polypeptide segment of herpes simplex virus (HSV) glycoprotein B1 or B2, useful in serodiagnostic immunoassays for distinguishing HSV-1 infection from HSV-2 in a human -
PS Claim 8; Fig 4; 23pp; English.
XX

The present invention provides antigenic peptides from herpes simplex virus type 1 (HSV1) and 2 (HSV2) glycoprotein B (gB) which can be used in the diagnosis of HSV infection, and identification of subtype, and in CC vaccines to protect against HSV. The present sequence is the HSV-1 gB protein.
CC Note: The present sequence is mentioned in claim 8 as being the HSV glycoprotein B1.
XX
SQ Sequence 904 AA;

Query Match 18.5%; Score 841; DB 22; Length 904;
Best Local Similarity 27.7%; Pred. No. 4.2e-53;
Matches 242; Conservative 179; Mismatches 345; Indels 108; Gaps 28;

Qy 35 PNTATWSTESPLTCHYGHDSHGE-RGNENRDSBEQNKNIYGPSFFPRVCSAGV 93
Db 80 PPRPA-----GDNATVAAGHATLREHLRDIKAENTDANFY-----VCPPTG 121
Qy 94 GDVFRFQTDHVCPDASD-MVHSEGLILYKQNIIPFMRVRKRYKVTTSTV----- 144
Db 122 ATVQFEQPRRCPTRPEGQNYTEGIAVFKENIAPYKFKATMYKDVTSQVWFHRSQ 181
Qy 145 YNGIYSDSITNOHTYKSIKPEW--TEKMDTIYQCFNSLRNLNTGNNLTYYVDRDDINMTV 202
Db 182 FMGIFED-----RAPVPEEVIDKINAKGVCRSTAKYVRNLETTFATFRDDHETDM 232
Qy 203 FLQPVGDV-----TPDVKRYGQPELYLPGHFWGSYRRRTTVNCELMDMFARSNPP 254
Db 233 ELKPANAARTSRGWHHTDLKNPGRVEAF-----HRYGTTVNCIVEEDARSYYP 283
Qy 255 FDFFTATGDTVMSPFWS-GEDDH-ENKMEKPFVSVINNY-KVDYQNRGTVPGLGKT 311

Db 284 YDEFVLATGDVYMSPPFYREGSHTERTSVAADRFXQVDFYARDLTTKARATAP--TT 341
Qy 312 RIFLDREBYTLSEKHLKMSYCPCLTLWKAPYNGIOTHSYHVFANDITASFTT---- 367
Db 342 RNLTTTPKFTVANDWVPRPSVCTMTKQVEDEMLRSEYGGSRFPSSDAISTTFTTNUTE 401
Qy 368 ---SKEDMKBFNTTYHCLNEEIKAEIKKYA-KVNSTHSKYGDILKYFKTDCGLVLMQPL 423
Db 402 YPLSRVDLGD-----CIGKDARDAMDRI FARRYNATHIKVGQPYLLANGGLFIAYQPL 455
Qy 424 IONRLLDKANKLNNTYSRRSRQAESTTDDPMEMTNGAGGEYSSENSITVAQVQYAYD 483
Db 456 LSNLTAE---LYVREHLREQSRKPPNPTPPPGASANASVERIKTTSSIEFARLQFTYN 511
Qy 484 NLRIRINNILEDLSKAWCREOHRAALVWNELSKINPTSVMSMIYNRPVSARKIGDVISVS 543
Db 512 HIQRHVNMDLGRVAIAWCEIQNHBLTLWNEARKLNPAIASATVGRRRVSARMGLGDMAVS 571
Qy 544 NCIVVDOTSVSLHKSRLLSASDEKCFSRPPVTTFKFMNDSTIYKQGLQVNNELITLTYL 603
Db 572 TCVPAADNVIVQNSMR-ISSRPGACYSRPLVSFRYEDOGPLVVGQGLGENNELRLTRDAI 630
Qy 604 ETCQENTEYFQAQKDMYIYKNEYHLKTVPLUSSITTLDTFTIALNFTLLENVDFKVIELYT 663
Db 631 EPCVTGHRRYFTFGGYVYFEEYAYSHQLSRADITTVSTFDLNTITMLEDFEFPVLEVYT 690
Qy 664 RDEKLSNVFDIETMFREYNYAQRVSLGRKDLDDLSTNRNQF--VDFAFGSMDLGLG-AV 720
Db 691 RHEIKDGLDYTEVQRNQLHDLRFADI-DTVIHADANAAMFAGLGAFFEGMGDLGRAV 749
Qy 721 GQTVNAVSGVATLFSISITGFINFIKNPFGGLMIIWIGVFAIYFTLTKTKIYETAP 780
Db 750 GKVMGIUGGVV----SAVSGVSSFSMNPFGALAVGLLVLAGLAAAPFAFRYVMRLQSNP 805
Qy 781 IKMIYPEIDK-LK-----EREKSEIAPISE-EELERIVLAMHIHQONSHMETK 827
Db 806 MKALYPLTTKELKUNTPNDPDSAGEGEGDFDEAKLAAREMIRYMALVSAMERTEH---K 862
Qy 828 TRKDPKDSILT-RAQNML---RRKSGYSNLKNAE 857
Db 863 AKKKGTSALLSAKAVTDMVRKERNNTNYTQVPNKD 896

RESULT 21
AAW72113
ID AAW72113 standard; Protein; 854 AA.
XX
AC AAW72113;
XX
DT 18-DEC-1998 (first entry)
XX
DE HSV-2 strain SB5 Contig ID 14 ORF#1 protein.
XX
KW HSV-2 strain SB5; immunological response induction; therapy;
KW antiviral identification; viral protein inhibitor.
XX
OS Herpes simplex virus type 2.
XX
PN WO9820016-A1.
XX
PD 14-MAY-1998.
XX
PF 31-OCT-1997; 97WO-US20016.
XX
PR 09-JUN-1997; 97US-0049018.
PR 04-NOV-1996; 96US-0030279.
XX
PA (SMIK) SMITHKLINE BEECHAM CORP.
XX
PI Chan JY, Dabrowski-Amaral CE, Delvecchio AM, Dillon SB;
PI Esser KM, Leary JJ;
XX

Db	76	TPPPD-----ANATVAAGHATLR--AHLREIKVENADAQ-----FVVCPPPT	115
Qy	93	VGDVFRFQTDHVCPSDASD-MVHSEGIILYKONIIPFMRVRKYRVKVTSTV-----	144
Db	116	GATVVOFEQPRRCPTREPGQNTYEGIAVFKENIAPYKATMYKDVTSQVWFGRHYS	175
Qy	145	-YNGIYSDSITNOHTFYKIEPWE--TEKMDTYOCFNSLRNTGNNLLTYVDRDDNMT	201
Db	176	QFMGIFED-----RAPVFEEVIDKINAKGVCRCSTAKYVRNNMETTAFHRDDHETD	226
Qy	202	VFLQPDVGVTPDKVYKSGOPELILPGFWGYSRRRTTVNCELMDFARSNPPDFVTA	261
Db	227	MELKPAK-VATRTSRGWHTTDLKYNPSRVEAFHYGTTVNCIVEEDARSVPYDFEFLA	285
Qy	262	TGDTVEMSPFWG-GEDDH-ENKMHKPFVFSVINNY-KVVDYQNRGTVPGLKTRIFLDRE	318
Db	286	TGDFVMSPFYGYREGSHTHTSYAADRFKQVDFYARDLTTKARATSP--TTRNLLTTP	343
Qy	319	EYTLSEKHLKNNMSYCPPLTWKAFYNGIQTEHSGSYHFVANDITASFTT-----SKED	371
Db	344	KFTVAMDVVPKPAVCTMTKQEVDEMRLAEYGGSFSSDAISTTFTTNLTQVSLSRVD	403
Qy	372	MKEFNTTYHCLNEEIKAEIEKYYA-KVNSTHSGYDGLKFKTDGGLYLWQPLIQNRLLD	430
Db	404	LGD-----CIGRDAREALDRMFARKYNATHIKVGQPOYYLATGGLFIAYQPLLSNTLAE	457
Qy	431	AKNNLNNETYSR-RSRROAESTTDPMMEM-TGNGAGGEYSSENSITVAQOVAYDNLIR	488
Db	458	---LVREYWEQDRKPNATPAPLREAPSANASVERIKTTSIEFARLQFTYNHIORH	513
Qy	489	INILEDLSKAWCRQHRALVWNLSEKINPTSPVMSIYNRPVSAKRIGDVISNCTIW	548
Db	514	VNDMLGRIAVAMCELOHQLHETLWNEARKLNPNAIASATVGRRSARMGLGVMAVSTCVPV	573
Qy	549	DOTSVLSHLSRLLSASDEKESRPPVTFKPMNDSTIYKQGLGVNNEILLTTTYLETQOE	608
Db	574	APDNVIVQNSMR-VSSRPGTCYSRLVSRFVYEDQOGLIEGOLGENNELRLTRDALEPCTV	632
Qy	609	NTEYVFOAKTDMYIKYNEHLKTVPLSSITITLDTFIALNFTLLENVDKVELYTRDEKR	668
Db	633	GHRRFIFGGGVVFEYAYSHQLSRADVTTVSTFIDLNITWLEDHFVPLEVYTRHEIK	692
Qy	669	LSNVFDIETMFREYNYAQRVSGRLKDLIDTSTNRNQFVD--AFGSLMDDLQ-AVGQTV	725
Db	693	DSGLLDYTEVQRRNGLHLRPAID-DTVIRADANAAMFAGLCAPEGMDLGRAVGKVM	751
Qy	726	NAVSVATLFSIYVTFINFPKPGMLIIVIGVLFAIYPLTKTKIYETAPIKIY	785
Db	752	GVVGWV----SAVSGVSSFSMNPFGALAVGLLVLAGLVAFAFPYVYLQLRNPMKALY	807
Qy	786	PEIDKLKEREGKSEIAPISSE-----ELERIVLAMHIHQONSHMETKTRKDP	832
Db	808	PLTTKELKTSPPGGVGGEGEAGGEGGDEAKLAAREMIRYMALUSAMETERHARKKG	867
Qy	833	KDSIL-TRAQNW-LRKR--SGYSNLKNAE	857
Db	868	TSALLSSKVTNNVLRKRKARYSPLHNE	896
RESULT 23			
AAW00376			
ID	AAW00376	standard; Protein; 904 AA.	
XX			
AC	AAW00376;		
XX			
DT	24-JAN-1997	(first entry)	
XX			
DE	HSV-2 glycoprotein B.		
XX			
KW	HSV-2; glycoprotein B; vaccine; diagnosis; immunoassay.		
XX			
OS	Herpes simplex virus type 2.		
XX			

FH Key Location/Qualifiers
 FT Peptide 1..22
 FT /label= Sig_peptide
 FT Peptide 18..75
 FT /label= Dominant_type-specific_epitope
 FT Domain 23..744
 FT /label= Extracellular_domain
 FT Domain 745..798
 FT /label= Transmembrane_domain
 FT Domain 799..804
 FT /label= Cytoplasmic_domain
 FT Peptide 819..904
 FT /label= Cross-reactive_antigenic_epitope
 XX
 PN W09632962-A1.
 XX
 PD 24-OCT-1996.
 XX
 PF 19-APR-1996; 96WO-US05316.
 XX
 PR 21-APR-1995; 95US-0426604.
 XX
 PA (UYNE-) UNIV NEW MEXICO STATE.
 XX
 PI Bell R, Goade DE, Jenison S;
 XX
 XX WPT; 1996-485557/48.
 XX
 PT New type-specific and cross-reactive Herpes Simplex Virus epitope(s)
 PT - used for diagnosing HSV-1 and/or HSV-2 infections and in vaccines
 PT for producing neutralising antibodies
 XX
 PS Disclosure; Fig 5; 37pp; English.
 XX
 CC Herpes simplex virus type 2 (HSV-2) glycoprotein B (gB2) (AAW00376)
 CC and HSV-1 gB1 (AAW00375) include type-specific and cross-reactive
 CC antigenic epitopes. Recombinant type-specific epitopes can be used
 CC in immunoassays to distinguish between HSV-1 and HSV-2 infections.
 CC Recombinant cross-reactive epitopes are useful for the diagnosis of
 CC both HSV-1 and HSV-2 infections. Both types of epitope can be used
 CC in vaccines.
 XX
 SQ Sequence 904 AA;

Query Match 18.5%; Score 839; DB 17; Length 904;
 Best Local Similarity 28.3%; Pred. No. 5.9e-53;
 Matches 246; Conservative 164; Mismatches 367; Indels 92; Gaps 27;

Qy	33	TSPPNTATWSTESPLTGHYTHDSSHGCRGNNENRDEEQNKNIYGSPTFFYRVCSASG	92
Db	76	TPPPD-----ANATVAAGHATLR--AHLREIKVENADAQ-----FVVCPPPT	115
Qy	93	VGDVFRFQTDHVCPSDASD-MVHSEGIILYKONIIPFMRVRKYRVKVTSTV-----	144
Db	116	GATVVOFEQPRRCPTREPGQNTYEGIAVFKENIAPYKATMYKDVTSQVWFGRHYS	175
Qy	145	-YNGIYSDSITNOHTFYKIEPWE--TEKMDTYOCFNSLRNTGNNLLTYVDRDDNMT	201
Db	176	QFMGIFED-----RAPVFEEVIDKINAKGVCRCSTAKYVRNNMETTAFHRDDHETD	226
Qy	202	VFLQPDVGVTPDKVYKSGOPELILPGFWGYSRRRTTVNCELMDFARSNPPDFVTA	261
Db	227	MELKPAK-VATRTSRGWHTTDLKYNPSRVEAFHYGTTVNCIVEEDARSVPYDFEFLA	285
Qy	262	TGDTVEMSPFWG-GEDDH-ENKMHKPFVFSVINNY-KVVDYQNRGTVPGLKTRIFLDRE	318
Db	286	TGDFVMSPFYGYREGSHTHTSYAADRFKQVDFYARDLTTKARATSP--TTRNLLTTP	343
Qy	319	EYTLSEKHLKNNMSYCPPLTWKAFYNGIQTEHSGSYHFVANDITASFTT-----SKED	371
Db	344	KFTVAMDVVPKPAVCTMTKQEVDEMRLAEYGGSFSSDAISTTFTTNLTQVSLSRVD	403
Qy	372	MKEFNTTYHCLNEEIKAEIEKYYA-KVNSTHSGYDGLKFKTDGGLYLWQPLIQNRLLD	430

```
Db 404 LGD-----CIGRDAREALDRMFARKYNATHIKVGQPOYLLATGFLIAYQPLSNTLAE 457
Qy 431 AKNKLNETYSR-RSRQAEISTTPMMEM-TGNAGGEYSSENSITVAQVQYADNLRIR 488
Db 458 ----LYVREYMRQDRKPRNATPAPLRAPASANASVERIKTTSIEFARLOFTYNNHIQRH 513
Qy 489 INNILEDLSKAWCREQHRALVWNLKINPTSVMSMIYNRPVSAGKRGDVISVNCIVV 548
Db 514 VNDMLGRIAVAWCELQNHLETLWNEARKLNPNATASATVGRVRSARMLGDVMAVSTCPV 573
Qy 549 DQTSVLHLSRLLSASDEKCFSPPTVTFKFMNDSTIYKGLGVNNEILLTTTYLETQCE 608
Db 574 APDNVIVQNSMR-VSSRPGTCYSRPLVSFRYEDQGLIEGOLGENNELRLTRDALEPCTV 632
Qy 609 NTEYFQAKTDMYIKNYEHLKTVPLSSITLDTFTIALNPTLLENVDFKVIELYTRDEKR 668
Db 633 GHRRYFIFGGGVYFEEYAYSHQLSRADVTTVSTFDLNTMLEDHFEFVPLEVYTRHEIK 692
Qy 669 LSNVFDIETMFREYNYAQRVSGRLKDLLSLTRNQFVD--AFGSLMDDLQ-AVGQTVV 725
Db 693 DSGLLDYTEVQRNQLHDLRFADI-DTVIRADANAAMPAGLCAPFEGMGDLGRAVGKVM 751
Qy 726 NAVSGVATLFSISITVTGFINFKNPFPGGMLMIIVVIGVLFALFYFLTKTKIYETAPIKMIY 785
Db 752 GVUGGVV----SAVSGVSSFSMSPFGALVGLLVLAGLVAAFPAPRYVLQLRNPKALY 807
Qy 786 PEIDKLEREGKSEIAPISBE-----ELERIVLAMIHQONSHMETKTRKOP 832
Db 808 PLTTKELKTSDPGGVGGEAGEGGGFDEAKLAAREMIRYMALVSAMERTEHKARKKG 867
Qy 833 KDSLIL-TQAQNM-LRKR--SGYSNLKNAE 857
Db 868 TSALLSSKVTNVLKRNKARYSPHNED 896

RESULT 24
AAW72193
ID AAW72193 standard; Protein; 904 AA.
XX
AC AAW72193;
XX
DT 13-JAN-1999 (first entry)
XX
DE HSV-2 strain SB5 Contig ID 15 ORF#28 protein.
XX
KW HSV-2 strain SB5; immunological response induction; therapy;
KW antiviral identification; viral protein inhibitor.
XX
OS Herpes simplex virus type 2.
XX
PN W09820016-A1.
XX
PD 14-MAY-1998.
XX
PF 31-OCT-1997; 97MO-US20016.
XX
PR 09-JUN-1997; 97US-0049018.
PR 04-NOV-1996; 96US-0030279.
XX
PA (SMIK ) SMITHKLINE BEECHAM CORP.
XX
PI Chan JY, Dabrowski-Amaral CE, Delvecchio AM, Dillon SB;
PI Esser KM, Leary JJ;
XX
DR WPI; 1998-286847/25.
DR N-PSDB; AAV62176.
XX
PT Herpes simplex virus type-2 sequences - useful in, e.g. prevention
PT and treatment of infection or inducing immunological response in
XX mammal
PS Claim 10; Page 121; 748pp; English.
```

```
XX This sequence represents a Herpes simplex virus type-2 (HSV-2) protein
CC sequence of the invention. This sequence was isolated from a HSV-2 strain
CC SB5 (deposited as ATCC VR-2546) DNA fragment designated Contig ID 15.
CC Based on homology, this sequence is a glycoprotein B precursor.
CC The proteins can be used for the treatment or prevention of disease, to
CC induce an immunological response in a mammal or to identify inhibitors,
CC activators or novel antivirals. Antagonists of the proteins can be used
CC to inhibit a viral polypeptide. The DNA sequence or a vector containing
CC it can also be used to induce an immunological response in a mammal.
XX
SQ Sequence 904 AA;
Query Match 18.5%; Score 839; DB 19; Length 904;
Best Local Similarity 28.3%; Pred. No. 5.9e-53;
Matches 246; Conservative 164; Mismatches 367; Indels 92; Gaps 27;
Qy 33 TSPNTATWSTESPLTGHYGTDSHGERGNENRDSSEQKNKNIYGSPTFPYRVCSASG 92
Db 76 TTPPD-----ANATVAAGHATLR--AHLREIKVENADAQ-----FVCCPPT 115
Qy 93 VGVFRFQTDHVCDDASD-MVHSEGILLIYKQNIIPMFVRKYRKVVTTSTV----- 144
Db 116 GATVWQFEQRRCPTRPEQONYTEGIAVVFKENIAPYKFKATMYKDVTVSQVMFGHRS 175
Qy 145 -YNGIYSDSITNQHTFYKSIPEWE--TEKMDTIYQCFNSRLNTGGNLLTYVDRDDINMT 201
Db 176 QFMGIFED-----RAPVPEEVIDKINAKGCRSTAKYVRNNMETTAFHRDDHETD 226
Qy 202 VFQPPVDGVTDPVKRYGSOPELYLEPCWFGSYYRRRTTNCVCLMDMFARSNPPDFPVTA 261
Db 227 MELKPAK-VATRTSRGHHITDLKYNFSRVEAFHYGTGTVNCIVEEDARSVYDFVLA 285
Qy 262 TGDTEVMSPPWS--GEDDH--ENKMHKXPWFVSVNNY-KVVDYQNRGTVPGLKTRIFLDRE 318
Db 286 TGDVFMVSPFYGYREGSHTSHTSYAADRFKQVDFYARDLTTKARATSP--TTRNLLTTP 343
Qy 319 EYTLSEKHLKMSYGLTLWKAIFYNGIQTEHSGSYHFVANDITASPTT-----SKED 371
Db 344 KFTVAMDWPKRPACTMTKQWDEMLRAEYGGSPFRSDAISTTFTTNLTQVLSRVD 403
Qy 372 MKEFNTTYHCLNEEIKAEIEKYYA-KVNSTHSGYGLKYFKPTDGLGLYVWQPLQIRLLD 430
Db 404 LGD-----CIGRDAREALDRMFARKYNATHIKVGQPOYLLATGFLIAYQPLSNTLAE 457
Qy 431 AKNKLNETYSR-RSRQAEISTTPMMEM-TGNAGGEYSSENSITVAQVQYADNLRIR 488
Db 458 ----LYVREYMRQDRKPRNATPAPLRAPASANASVERIKTTSIEFARLOFTYNNHIQRH 513
Qy 489 INNILEDLSKAWCREQHRALVWNLKINPTSVMSMIYNRPVSAGKRGDVISVNCIVV 548
Db 514 VNDMLGRIAVAWCELQNHLETLWNEARKLNPNATASATVGRVRSARMLGDVMAVSTCPV 573
Qy 549 DQTSVLHLSRLLSASDEKCFSPPTVTFKFMNDSTIYKGLGVNNEILLTTTYLETQCE 608
Db 574 APDNVIVQNSMR-VSSRPGTCYSRPLVSFRYEDQGLIEGOLGENNELRLTRDALEPCTV 632
Qy 609 NTEYFQAKTDMYIKNYEHLKTVPLSSITLDTFTIALNPTLLENVDFKVIELYTRDEKR 668
Db 633 GHRRYFIFGGGVYFEEYAYSHQLSRADVTTVSTFDLNTMLEDHFEFVPLEVYTRHEIK 692
Qy 669 LSNVFDIETMFREYNYAQRVSGRLKDLLSLTRNQFVD--AFGSLMDDLQ-AVGQTVV 725
Db 693 DSGLLDYTEVQRNQLHDLRFADI-DTVIRADANAAMPAGLCAPFEGMGDLGRAVGKVM 751
Qy 726 NAVSGVATLFSISITVTGFINFKNPFPGGMLMIIVVIGVLFALFYFLTKTKIYETAPIKMIY 785
Db 752 GVUGGVV----SAVSGVSSFSMSPFGALVGLLVLAGLVAAFPAPRYVLQLRNPKALY 807
Qy 786 PEIDKLEREGKSEIAPISBE-----ELERIVLAMIHQONSHMETKTRKOP 832
Db 808 PLTTKELKTSDPGGVGGEAGEGGGFDEAKLAAREMIRYMALVSAMERTEHKARKKG 867
```

QY 833 KDSIL-TRAQNM-LRKR--SGYSNLKNAE 857
 Db 868 TSALLSSKVTNNVLKRKARYSPLHNE 896

RESULT 25
 AAB74442
 ID AAB74442 standard; protein; 904 AA.
 AC AAB74442;
 XX 29-MAY-2001 (first entry)
 DE Herpes simplex virus 2 glycoprotein B.
 KW HSV-1; HSV-2; glycoprotein B; gb; transmembrane envelope glycoprotein;
 KW antigenic epitope; diagnosis; vaccine.
 XX Herpes simplex virus type 2.
 OS US6197497-B1.
 PN 06-MAR-2001.
 PD 19-APR-1996; 96US-0632537.
 PR 21-APR-1995; 95US-0426604.
 PA (UYNE-) UNIV NEW MEXICO STATE.
 XX Goade DE, Bell R, Jenison S;
 PI WPI; 2001-256360/26.
 DR Continuous, isolated, antigenic polypeptide segment of herpes simplex
 PT virus (HSV) glycoprotein B1 or B2, useful in serodiagnostic
 PT immunoassays for distinguishing HSV-1 infection from HSV-2 in a human
 XX Claim 1; Fig 5; 23pp; English.
 CC The present invention provides antigenic peptides from herpes simplex
 CC virus type 1 (HSV1) and 2 (HSV2) glycoprotein B (gb) which can be used in
 CC the diagnosis of HSV infection, and identification of subtype, and in
 CC vaccines to protect against HSV. The present sequence is the HSV-2 gb
 CC protein.
 CC Note: The present sequence is mentioned in claim 8 as being the HSV
 CC glycoprotein B2.
 SQ Sequence 904 AA:

Query Match 18.5%; Score 839; DB 22; Length 904;
 Best Local Similarity 28.3%; Pred. No. 5.9e-53;
 Matches 246; Conservative 164; Mismatches 367; Indels 92; Gaps 27;

QY 33 TSPNTATWSTESPLTGHYTHDSHGERNNENRDSBEQNKIYGSPSTPPYRCSASG 92
 Db 76 TTPPD-ANATVAAGHATLR-AHLREIKVENADAQ-----FYVCPPT 115
 QY 93 VGDVFRQTDHVCPPASD-MVHSEGLILYKQNIIPFMRVRKRYKVVTTSTV----- 144
 Db 116 GATVVQFEQPRCPTREPGQNTGEGIAVVFKEKNIAPYKFKATMYKYKDVTSQVWFGHRY 175
 QY 145 -YNGIYSDSITNQHTFYKSIEPWE--TEKMDIYOCFNSLRNTGNNLLTYVDRDDINMT- 201
 Db 176 QMGIFED-----RAPVFEVIDKINAKGVCSTAKYVRNNMETTAFHRDDHETD 226
 QY 202 VFLQPVGDVTPDKRYGSOPELYLEPGFWGSGYRRRTTNCCLMDFMARSNPPDFEFTVTA 261
 Db 227 MELKPAK-VAITSTSGMHTDLKYNPSRVEAFHRYGTTVNCIVEVDARSVPDFEFLA 285
 QY 262 TGDVTVMSPFWS-GEEDH-ENKMHKPKFWFSVINNY-KVDYQNRGTVPVPLGKTRIFLDRE 318
 Db 286 TGDVYMSPPFYGYREGSHTENTSAAADRFPKQVDFYARDLTTKARATSP--TTRNLLTTP 343

319 EYTLSEKHLKMSYCPPLTLWKAFVNGIOTBSHSGSYHFVANDITASFTT-----SKED 371
 Db 344 KFTVANDWVPRPAVCTMTKQOEVDMLRABYSGSFRESSDAISTFTTNTLTQVLSLSD 403
 QY 372 MKEFTNTYHCLNEEIKABIEKKA-KVNSTHSKYGLKYFKTGDGLYLVMQPLQNRLLD 430
 Db 404 LGD-----CIGRDAREAIDRMFARKYNATHIKVGQPOYFLATGGFLIAYQPLLSNTLAE 457
 QY 431 AKNNLNNETYSR-RSRQAEESTTDPMMEM--TCNGAGGEYSSENSITVAQVAYDNLIR 488
 Db 458 ----LYVREYMRQDRKPRNATPAPLRAPASANASVERIKTTSSTIEFARLQFTYNHQRH 513
 QY 489 INNILEDLSKAWCREQHRALVWNELSKINPTVMSMIYNRPVSAKRIGDVISVNCIWW 548
 Db 514 VNDMLGRIAVAWCLQNHLELTWNEARKLNPNATASATVGRVSRMGLDVMVSTCVPV 573
 QY 549 DOTSVLSHLSRLLSASDEKCFSPPTVTFKPMNDSTIYKGLGVNNEILLTTTLYLETCOE 608
 Db 574 APDNVIVQNSMR-VSSRPGTCYSRPLVSFRYEDQGLIEGQLEGENNELRLTRDALEPCTV 632
 QY 609 NTEYFQAKTDMYIKNYEHLKTVPLSSIITLDTFIALNPTLLENVDFKVIETYRDEKR 668
 Db 633 GHRRYFIFGGGVYFEEYAYSHQLSRADVTTVTFIDLNITMLDEHFVPLEVYTRHEIK 692
 QY 669 LSNVFDIETMPREYNYAQRVSLRKLDDLSTNRNQFVD--AFGSLMDDLGL-AVGQTVV 725
 Db 693 DSGLLDYTEVORRNQLHDLRFADI-DTVIRADANAAMPAGLCAFFEGMCDLGRAVGKVM 751
 QY 726 NAVSGVATLFSSIVTGFINFINKPFGGMLMIIVVIGVLFAIYFLTKTKIYETAPIKMIY 785
 Db 752 GVVGGVV-----SAVSGVSSFMSPFGALAVGLLVLAGLVAFAFRYVVLQQRNPMKALY 807
 QY 786 PEIDKLKEREKSEIAPISEE-----ELERIVLAMHIHQNSHMETKTRKDP 832
 Db 808 PLTTKELKTSPPGGVGESEAGGFGDEAKLAAREMIRYMALVSAMERTEHKARKKG 867
 QY 833 KDSIL-TRAQNM-LRKR--SGYSNLKNAE 857
 Db 868 TSALLSSKVTNNVLKRKARYSPLHNE 896

RESULT 26
 AAP60244
 ID AAP60244 standard; protein; 903 AA.
 XX AC AAP60244;
 XX 25-MAR-2003 (updated)
 DT 31-JUL-1991 (first entry)
 DE Herpes simplex virus glycoprotein gb.
 KW HSV; gb glycoprotein; vaccine.
 OS Herpes simplex virus.
 XX EPI70169-A.
 XX 05-FEB-1986.
 PF 19-JUL-1985; 85EP-0109042.
 XX 20-JUL-1984; 84JP-0151766.
 PR 11-DEC-1984; 84JP-0262465.
 XX (KAGA) CHERO SERO THERAPEUTIC RES INST.
 XX Noraki C, Makizumi K, Kino Y, Eto T, Ohtomo N;
 WPI; 1986-036935/06.
 DR N-PSDB; AAN60195.
 XX


```

Db 170 --RAPVPE- EYWDKINAGVCKSTAKYVRNNMESTAFHRDDHESDMALKPAKAAAT-RTS 225
Qy 216 RYGSQPELYLEPGWFWGVRRTTTCVNCELMDFAFNSPPDFVATGDTVMSPFWGSE 275
Db 226 RGWHTTDLKYNPARVEAFHRYGTTTNCIVVEEYEAESVYPYDFVLATGDFVYMSPFYGR 285
Qy 276 D-DH-ENKMKHEKPFWSVINNY-KVVDYQNRGTVPGLKTRIFLDREEVYTLSEKHLKMS 332
Db 286 DSGHGEHTAYADRFRQVDGYERDLSTGRRAAPV--TRNLLTPKFTVGDWAPKPS 343
Qy 333 YCPLTLKAFYNGIQTEHSGSYHFVANDITASFTT-----SKEDMKEFNTTYHCLNEE 385
Db 344 VCTLTKWREDEMLRAEYGFSPFSSAALSTFTTANRTEYALSRLVDLAD-----CVGRE 397
Qy 386 IKAIEKKY-AKVNTHSKYGLKFKYKTDGGLYVWQPLIQNRLLDAKKNLNEYYSRS 444
Db 398 AREAVDRIFLRRYNGTHVKVGQVYLLATGGFLIAYQPLLSNALV-----ELYREL 449
Qy 445 RRAESTTDPMMEMTGNAGGE-----YSSENSITVAOVQYAYDNL 486
Db 450 VR-----EQTRPAGGDPGEAATPGSPVDPSPVERIKTSSVEFARLQFYDHIQ 499
Qy 487 TRINNLEDLSKAWCREQHRAALVWNELSKINPTSMGMIYNRPVSAKRIGDVISVNCI 546
Db 500 RHVNDMLGRITAWCELQNRLLTWNEARRLNPGAISATVGRVRSARMLGDVMAVSTCV 559
Qy 547 VVDQTSVLSHLKRLLSASDEKCFRPPVTFKPMNDSTIYKQGLVNNELLTITTYLET 606
Db 560 VPAPDNVIMQNSMR-VAARPQTCYGRPLVSFRYEAADGPLVEQLGEDNEIRLERDALBPC 618
Qy 607 QENTVEYFOAKDMYIKNYEHLKTVPLSSITTLDTFIALNFTLENVDFKVIELYTRDE 666
Db 619 TVGHRRYTFGAGYVFEYAYSHQGLRADVTIVTFINLNTMLDEHFEVPLEVYTRQE 678
Qy 667 KLSNVFDIETFREYNYAQRVSGRLKDLDDLSTNRNQVDAFGLSMDLDDGAVGQTVYN 726
Db 679 IKDSGLLDVTEVORRNQHLARFADI-DTVIKADAAHALFAGLY-SFFEGLDGUGRAVGK 736
Qy 727 AVSGVATLFSSIVTGFINKPFGMLMIIVVIGVLFAYIPLTKTKIYETAPIKMIYP 786
Db 737 VMGIVGVWSAVSGVSSFLSNPFGALAVGLLVLAGLAAAFPAFRYVRLQRPNMKALYP 796
Qy 787 EIDKLKREGKEIAPIS-----BEELERIVLA-----MHTHQNSHMETKTRKD 832
Db 797 ----LTKELSDGAPLAGGEGDGAEDFDEAKLAQAREMIRYMALVSAMERTEHKARKG 852
Qy 833 KDSIL---TRAQNMLKRSYGSLNKAESVEM 861
Db 853 TSALLSAKVTDAVMEKRAPRYSPLRDTDEEL 885

RESULT 28
AAW34552
ID AAW34552 standard; Protein; 904 AA.
AC AAW34552;
XX
DT 26-FEB-1998 (first entry)
DE Herpes simplex virus type 1 glycoprotein B.
KW Herpes simplex virus; HSV; type I; type II; glycoprotein B; vaccine;
KW membrane protein; virus-specific glycoprotein;
KW transmembrane anchor region.
OS Herpes simplex virus type 1.
FH Key Location/Qualifiers
FT Domain 726..795
FT Region /label= membrane spanning_region
FT /note="putative"

```

```

XX US5648079-A.
XX 15-JUL-1997.
XX 08-DEC-1994; 94US-0351875.
XX 20-OCT-1986; 86US-0921730.
XX 06-APR-1984; 84US-0597784.
XX 20-SEP-1990; 90US-0587179.
XX 21-DEC-1992; 92US-0993415.
XX 18-OCT-1993; 93US-0138717.
XX (CHIR ) CHIRON CORP.
XX Burke RL, Pacht C, Valenzuela PDT;
XX WPI; 1997-372022/34.
XX N-PSDB; AAT93650.
XX Vaccines against herpes simplex virus infection - containing
XX recombinant HSV glycoprotein B
XX Disclosure; Fig 4A-F; 33pp; English.
XX The present sequence is from the Herpes simplex virus (HSV) type I
XX strain Patton and represents a glycoprotein B. HSV is a double stranded
XX virus packaged within an icosahedral nucleocapsid enveloped within a
XX membrane. The membrane includes a number of virus-specific
XX glycoproteins, with glycoprotein B being one of the most abundant.
XX Glycoprotein B from both HSV type I and type II are cross reactive. New
XX vaccines comprising recombinantly produced glycosylated glycoprotein B
XX that has a deletion of at least part of the transmembrane anchor region,
XX in combination with a carrier and an adjuvant have been produced. The
XX vaccines are used for immunising humans against HSV (HSV type I or
XX HSV type 2) by vaccination before or after primary infection with HSV.
XX
SQ Sequence 904 AA;

```

Query Match 18.4%; Score 838; DB 18; Length 904;

Best Local Similarity 28.0%; Pred. No. 7e-53;

Matches 243; Conservative 166; Mismatches 368; Indels 92; Gaps 26;

```

Qy 33 TSPNTATWSTESPLTGHYTHDSSHGCRNENRDSSEQNKNIYGSPTFPRVCSASG 92
Db 76 TTPPD-----ANATVAAGHATLR--AHLREIKVENADAQ-----FYVCPPT 115
Qy 93 VGDVRFOTDHVCPDASD-MVHSEGLLIYKQNIIPPMFRVRKYRKYVTTSTV----- 144
Db 116 GATVVQEPQRCRTRPEGQNYTEGIAVFKENIAPYKFKATMYKDVTVSQVWFGHYS 175
Qy 145 -YNGIYSDSITNQHTFYKSIETPWE--TEKMDTIYOCFNSLRNLNTGGNLLTVYVDRDDINMT 201
Db 176 QFMGIFED-----RAPVPFEVIDKINAKGVCSTAKYVRNNMETTAFHRDDHETD 226
Qy 202 VFLOPVDGVTDPVKYGSQPELYLEPGWFWGYSYRRRTVNCELMDFAFNSPPDFVTA 261
Db 227 MELAPAK-VATRTSRGWHTTDLKYNPSRVEAFHRYGTTVNCIVEEDARSVYPYDFVLA 285
Qy 262 TGDVEMSPFWS-CEDDH-ENKMKHEKPFWSVINNY-KVVDYQNRGTVPGLKTRIFLDRE 318
Db 286 TGDVYMSPFYGYREGSHTSHTSVAAADRKFQVDFYARDLTTKARATSP--TTRNLLTTP 343
Qy 319 EYTLSEKHLKMSYCPCLTLWKAIFYNGIQTEHSGSYHFVANDITASFTT-----SKED 371
Db 344 KFTVADWVPRPAVCTMTKQVEDEMLRAEYGFSPFSSAALSTFTTANRTEYALSRLVD 403
Qy 372 MKEFTTYHCLNEEIKAEIEKKYA-KVNSTHSGYGLKFKYKTDGGLYVWQPLIQNRLLD 430
Db 404 LGD-----CIGRDAREAIDRMFARKYNATHIKVGQFYLLATGGFLIAYQPLLSNTUAE 457
Qy 431 AKNKLNNETYSR-RSRRAESTTDPMMEM-TGNCAGGEGYSSENSITVAOVQYAYDNLRI 488

```

Db 458 ----LYVREYVREQDKPRNATPAPLREAPGASANASVERIKTSSIEFARLQFTYNNHQH 513
 Qy 489 INNILEDLSKAWCREQHRALVWNELSKINPTSVMSIYNRPVSAKRIGDVISVNCIVV 548
 Db 514 VNDMLGRIAVAWCBLQNHLETLWNEARKLNPNATASATVGRVRSARMLGDVMAVSTCPV 573
 Qy 549 DOTSVLSHLSRLLSASDEKCFSPRPVTFKFMNDSTIYKGLGVNNEILLTITTYLETQCE 608
 Db 574 APDNVIVQNSMR-VSSRPGTCYSRPLVSFRYEDQGPLIEGQGENNELRLTRDALEPCTV 632
 Qy 609 NTEYFFQAKTDMYIKYKYEHLKTVPLSSITTLDTFIALNFTLLENVDFKVIELVTRDEKR 668
 Db 633 GHRRYFIFGGGVYVFEYAYSHQLSRADVTITVSTFIDLNITMLDHEBFVPLEVYTRHEIK 692
 Qy 669 LSNVFDIETPREYNYAQRVSGRLKDLLDLSLTVRNOFVD--AFGSLMDDLQ-AVGQTVV 725
 Db 693 DSGLLDTEVORRNQLHDLRFADI-DTIVRADANAAMFAGLCAFEFGMGDLGRAVGKVM 751
 Qy 726 NAVSGVATLFSIIVTGFINFIKNPFGGMMLIIVVIGLVFAIYFLTKTKIYETAPIKMIY 785
 Db 752 GWVGWV---SAVSGVSSFMSPFGALAVGLLVAGLVAAFFAFRYVLQLRNPMKALY 807
 Qy 786 PEIDKLKERECKSIAPISE-----ELERIVLAMHIHOONSHMETKTRKDP 832
 Db 808 PLTTKELKTSDPGGVGGEAGEGEGGFDEAKLAAREMIRYMALVSAMERTEHKARKKG 867
 Qy 833 KDSIL-TRAQNM---RKRSGYSLKNAE 857
 Db 868 TSALLSSKVTNVLAKNKARYSPLHNE 896

RESULT 29

AAP711136
 ID AAP711136 standard; Protein; 907 AA.

AC AAP711136;

CC 25-MAR-2003 (updated)

DT 30-APR-1991 (first entry)

XX Herpes Simplex Virus-2 gB from p52BXX.

XX Vaccine; prophylaxis; treatment; Herpes Simplex Virus-1; glycoprotein; gB.

OS Herpes simplex virus type 2 HG52.

XX US4642333-A.

XX 10-FEB-1987.

XX 20-JUN-1984; 84US-0622496.

XX 20-JUN-1984; 84US-0622496.

XX (PERS/) PERSON S.

XX Person S;

XX WPI; 1987-056354/08.

XX N-PSDB; AAN71399.

XX Amino acid chain of glyco:protein B of HSV-1 and 2 - prepd. as
 PT recombinant and used for vaccines for herpes simplex virus types 1
 PT and 2.

XX Example; Table 2; 16pp; English.

XX A pure non-glycosylated amino acid (AA) chain comprising a sequence
 CC corresponding to that occurring in glycoprotein B of HSV-1 or HSV-2
 CC virus, which is antigenic to HSV-1 of HSV-2, which contains no more
 CC than 750 AA residues, and which includes AA residues 135-649
 CC inclusive is claimed. It can be used to produce vaccines for

CC prophylaxis and treatment of HSV-1 and HSV-2.
 CC See AAN71303 for the HSV-1 sequence.
 CC (Updated on 25-MAR-2003 to correct PR field.)
 XX Sequence 907 AA;
 SQ
 Query Match 18.4%; Score 838; DB 8; Length 907;
 Best Local Similarity 28.8%; Pred. No. 7e-53;
 Matches 241; Conservative 163; Mismatches 347; Indels 86; Gaps 26;
 Qy 65 ENRDSBQNKNIYGSPTFPYRVCSAGVGDFRQTDHVCPPDASD-MVHSEGLLIYKQ 123
 Db 105 ENADAQ-----FYVCPPTGATVVOFEQPRECPTRPEQONTEGIAVVFKE 150
 Qy 124 NIIPFVRVRYKVKVTTSTV-----YNGIYSDSITNQHTFYKSIKIEWE--TEKMDT 173
 Db 151 NIAPYKFKATMYKDYKTVTSQVWFGRHYSQFMGIFED-----RAPVFEYVIDKINA 201
 Qy 174 IYOCFNSRLRLNTGNNLLTVYDRODINNVTFLQPDVGTDPVKRYGSOPELYLEPGWFGS 233
 Db 202 KGYCRSTAKVVRNMETAFHRDDHETDMELKPAK-VATRTSGWHTTDLKYNPSRVEAF 260
 Qy 234 YRRRTTNCMLDMFARSNPFPDFVTATGDTVMSPFWS-GEDDH-ENKMHKPFVSV 291
 Db 261 HRYGTTVNCIVEVDARSVVPYDEFVLATGDFVMSPPFYCYREGSHTEHTTYAADRFKQV 320
 Qy 292 INNY-KVVDYQNRGTVPLGKTRIFLDREYTLSEKHLKNMSYCPLTLWKAFYNGIOTEH 350
 Db 321 DGFYARDLTTKARATAP--TTRNVLTTPKFTVAMDVPRKPAVCTMTKQEVDEMRAEY 378
 Qy 351 SGSYHFVANDITASFTT-----SKEDMKFEPTTYHCLNEEIKABIEKKYA-KVNSTHS 402
 Db 379 GGSRESSDAISTFTTNTLSEYSLRVDLGD-----CIGRDAREALDRMFARKYNATHI 432
 Qy 403 KYGDLKYFTDGGLYLWQPLIQNRLLDAKNLNNETYSR-RSRQAEESTTDPMMEM-TG 460
 Db 433 KVGOPYQATGGFLIAYQPLLSNTLAE---LYVREYMRQEQSRKPGNATPAFLREAPSA 488
 Qy 461 NGAGGEYSSENSITVAQOVAYDNLIRINNTILEDLSKAWCREQHRALVWNELSKINPT 520
 Db 489 NASVERIKTSSIEFARLQFTYNNHQHVNMDLGRIVAWCELQNHLETLWNEARKLNPN 548
 Qy 521 SVMSIYNRPVSAKRIGDVISVNCIWDQTSVLSHLRLLSASDEKCFSPRPVTFKPM 580
 Db 549 ATASATVGRVRSARMLGDVMAVATCPVAPDNVIVQNSMR-VSSRPGTCYSRPLVSFRYE 607
 Qy 581 NDSTIYKGLGVNNEILLTITTYLETQCENTEYFQAKTDMYIKYKYEHLKTVPLSSITTL 640
 Db 608 DQGPLIEGQGENNELRLTRDALEPCTVGHRRYFIFGGGVYVFEYAYSHQLSRADVTTV 667
 Qy 641 DTFIALNFTLLENVDFKVIELVTRDEKRLSNVFDIETMFREYNYAQRVSGRLKDLLDLS 700
 Db 668 STFIDLNITMLDHEBFVPLEVYTRHEIKDSGLLDYTEVQRRNQLHDLRFADI-DTIVRAD 726
 Qy 701 TNRNQFVD--AFGSLMDDLQ-AVGQTVVNAVSQVATLFSIIVTGFINFIKNPFGGMMLI 757
 Db 727 ANAAMFAGLCAFEFGMGDLGRAVGKVMGVGVV----SAVSGVSSFMSPFGALAVGL 782
 Qy 758 VVIGLVFAIYFLTKTKIYETAPIKMIYPEIDK-LK-----EREGKSEIAPISEEE 807
 Db 783 LVLAGLVAFAFFRYVLQLRNPMKALYPLTTKELKLPDPGGVGGEAGEGEGGFDEAK 842
 Qy 808 L---ERIVLAMHIHOONSHMETKTRKDPKDSIL-TRAQNM---RKRSGYSLKNAE 857
 Db 843 LAAREMIRYMALVSAMERTEHKARKKGTSSALLSSKVTNVMVRKNRKARYSPLHNE 899

RESULT 30

AAR41779
 ID AAR41779 standard; Protein; 904 AA.

XX AAR41779;

XX

DT	25-MAR-2003 (updated)	Db	220	TTAFHRDDHETDMELKPAANAATRTSRGWHHTDLKYNFSRVEAF	-----HRYGTTVN	270
DT	29-MAR-1994 (first entry)	Qy	242	CELWDMFARSNPDPDFVTATGDTVEMSPFWS-GEDDH-ENKMEKMPFVSVINNY-KVV	298	
DE		Db	271	CIVEVDARSVPYDFEFLATGDFVYMSPPFYGYREGSHTEHTSYAARDKQVDGYFARDL	330	
XX		Qy	299	DYQNRGTVPGLKTRIFLDREBYTLSEKHLKNMSYCYPLTLWKAFYNGIQTHSHSGSYHFA	358	
KW	Glycoprotein B (gB2).	Db	331	TTKARATAP--TTRNLLTTPKFTVAMDVWPKRPSVCTMTWKQEVDEMLRSEYSGSFRSS	388	
KW	Glycoprotein; gB1; gB2; herpes simplex virus, HSV-1; HSV-2; flanking; initiation; termination; transcription; translation; regulatory sequence.	Qy	359	NDITASFTT-----SKDMKEFTNTYHCLNEBIKAEIEKKYA-KVNSTHSKYGDLKYP	410	
XX		Db	389	DAISTTTTNTLTPELSRVDLGD-----CIGKQDAMDRIFARRVNATHIKVGQPOYY	442	
OS	Herpes simplex virus 2 strain 333.	Qy	411	KTDGGLVLMQPLIONRLLDKAKNKNNETYSRRSRRAESTTDDPMEMTNGAGGEYSSE	470	
XX		Db	443	LANGGFLIAYOPLLSNLTAE-----LYVREHUREOSRPPNPTPPPPGASANASVERIKIT	498	
FT	Key Location/Qualifiers	Qy	471	NSITVAQVQAYDNLIRINNILEDLSKAWCREQRAALVWNELSKINPTSVMSMYNRP	530	
FT	Misc-difference 82 /note= "ARG encoded by GCG (sic)"	Db	499	SSIEFARLQFTYNIQRHVNDMLGRVAIAWCELQNHETLWNEARKLNPAIASATVGR	558	
FT	Misc-difference 478 /note= "ASN encoded by ACC (sic)"	Qy	531	VSAKRIGDIVSNCIVVDQTSVSLHKSLLSASDEKCFSPRPVTFKFMNDSTIYKQOL	590	
FT	Misc-difference 540 /note= "ASN encoded by ACC (sic)"	Db	559	VSARMLGDVMAVSTCVPAADNVIVQNSMR-ISSRPGACYSRPLVSFRYEDQGPLVEGOL	617	
FT	Misc-difference 882 /note= "ARG encoded by GCC (sic)"	Qy	591	GVNNEILLTTLTETCQENTEYEFQAKTDMYIKNYEHLKTVPLSSITLDTFIALNFTL	650	
XX	US5244792-A.	Db	618	GENNELRLTRDAIEPCTVGHRRYFTFGGYVYFESAYSQHLSRADITTVSTFIDLNITM	677	
XX	14-SEP-1993.	Qy	651	LENVDKVIELYTRDEKRLSNVFDIETMFREYNYAQRVSGRLRDLKDLNTRNQF--VD	708	
XX	20-SEP-1990; 90US-0587179.	Db	678	LEDHEFVPLEVYTRHEIKDSGLLDYTEVQRNQLDLRFADI-DTVIHADANAAMFAGLG	736	
XX	06-APR-1984; 84US-0597784.	Qy	709	AFGLMDDLG-AVQGTVMNAVSGVATLFSSIVTGFINFIRKPPFGMLMIIVIGVLFAYI	767	
PR	20-OCT-1986; 86US-0921730.	Db	737	AFPEGMDLGRAVGKVMGIVGGVV-----SAVSGVSSFMSPFGALAVGLLVLAGLAAAF	792	
XX	20-SEP-1990; 90US-0587179.	Qy	768	FLTKTKTIYETAPIKMIYPEIDK-LK-----ERECKSEIAPISE-EELERIVLA	814	
PA	(CHIR) CHIRON CORP.	Db	793	PAFYVMRLQSNPMKALYPLTTRELKNPTNPDSGEGEGEDFDEAKLAAREMIRYNAL	852	
PI	Burke RL, Pechl C, Valenzuela PDT;	Qy	815	MHIHQNSHMETKTRKDPKDSILT-RAQNML---RKESGYSNLKNAE	857	
XX	WPI; 1993-302641/38.	Db	853	VSAMERTEH---KAKKGTSAALLSAKVTDMWRKRNNTNYTOVENKD	896	
DR	N-PSDB; AAQ48496.	XX				
PT	DNA construct for expressing HSV glycosylated polypeptide -					
XX	useful for vaccinating against HSV-1 and -2 infections					
PS	Disclosure; Fig 4; 33pp; English.					
CC	DNA constructs for expressing a glycosylated polypeptide in a					
CC	eukaryotic host cell comprises (1) an oligonucleotide sequence (OS)					
CC	free of natural flanking sequences, encoding glycoprotein B (gB)					
CC	of HSV, or its fragments, and (2) 5' initiating and 3' terminating					
CC	transcriptional and translational regulatory sequences flanking OS,					
CC	at least one of these sequences not being from HSV. The HSV-1 and					
CC	HSV-2 sequences are given in AAQ48496 and AAQ48497 respectively.					
CC	(Updated on 25-MAR-2003 to correct PF field.)					
XX						
SQ	Sequence 904 AA;					
Query Match	18.4%; Score 836.5; DB 14; Length 904;					
Best Local Similarity	27.5%; Pred. No. 9e-53;					
Matches 244;	Conservative 181; Mismatches 361; Indels 101; Gaps 28;					
Qy	24 SLSIATGTGTPPNTATWSTESP--LTGHYGTHTDSSHGE-RGNNEENRDESEQNKNYIGSP	80				
Db	58 ALGAAPTGDPKPKKNKPNPTPPRPGDNATVAAGHATLREHLRDKAENTDANFY---	114				
Qy	81 STFFPRVCSAGVGDFRFTDTHVCPDASD-MVHSEGLLLYKQNIIPFMRVRKYRW	139				
Db	115 -----VCPPTGATVQVFEQPRRCPTREPQGNYTEGIAVVFKEINIAFYKFKATMYIKDV	168				
Qy	140 TTSTV-----YNGIYSDSITNQHFTFYKSTEPWE--TEKMDTIYQCFNSLRNLGTGNL	189				
Db	169 TVSQWFGHRYSQFGIIFED-----RAPVPEEVIDKINAKGVCRCSTAKYVRNMLE	219				
Qy	190 LTYVDRDDINMTVFLQPDGV-----TPDKRYGSGPELYLEPGWFGWSYRRRTTVN	241				

```

PR 04-NOV-1996; 96US-0030279.
XX (SMIK ) SMITHLINE BEECHAM CORP.
XX Chan JY, Dabrowski-Amaral CE, Delvecchio AM, Dillon SB;
XX Esser KM, Leary JJ;
XX WPI; 1998-286847/25.
XX N-PSDB; AAV62146.
XX Herpes simplex virus type-2 sequences - useful in, e.g. prevention
XX and treatment of infection or inducing immunological response in
XX mammal
XX Claim 10; Page 64; 748pp; English.
XX
XX This sequence represents a Herpes simplex virus type-2 (HSV-2) protein
XX sequence of the invention. This sequence was isolated from a HSV-2 strain
XX SB5 (deposited as ATCC VR-2546) DNA fragment designated Contig ID 91.
XX Based on homology, this sequence is a glycoprotein B precursor.
XX The proteins can be used for the treatment or prevention of disease, to
XX induce an immunological response in a mammal or to identify inhibitors,
XX activators or novel antivirals. Antagonists of the proteins can be used
XX to inhibit a viral polypeptide. The DNA sequence or a vector containing
XX it can also be used to induce an immunological response in a mammal.
XX
XX Sequence 795 AA;
XX
XX Query Match 18.4%; Score 836; DB 19; Length 795;
XX Best Local Similarity 29.1%; Pred. No. 8.1e-53;
XX Matches 237; Conservative 157; Mismatches 349; Indels 72; Gaps 24;
XX
XX 87 VCSASGVDVFRFQDTHVCPDASD-MVHSEGLIYKQNIIPFMRVRYKRVKVVTTSTV- 144
XX 1 VCPPTGATVVFQFQRRCPTRPEGQNYTEGIAVVFKENIAPYKFKATMYKDVTSQVW 60
XX
XX 145 -----YNGIYSDSIHQHTFYKSIPEWE--TEKMDTIYQCFNSLRINTGGNLLTYVDR 195
XX 61 FGHYSQFMGIFED-----RAPVFEVIDIKNAKGVCRSTAKYVRNNMETTAFHR 111
XX
XX 196 DDINMTVLPQVDGTPDKRYGQPELYLEPGFWGYSRRRTTVCNCELMDFARSNPPF 255
XX 112 DDHETDMELPAK-VATRTSRGWHTTDLKYNPSRVEAFHRVGTIVNCIVEVDARSVYPY 170
XX
XX 256 DFFVTATGDTVEMSPFWS-GEEDH-ENKHEKPFVSVINNY-KVDYQNRGTVPVLKTR 312
XX 171 DEFVLATGDFVYMGFFYGYREGSHTSTSYAADRFRQVDGFYARDLTTKARATSP--TTR 228
XX
XX 313 IFLDREEVTLSEWEKHLKNMSYCPLLTWKAFYNGIQTEHSGSYHFVANDITASFTT----- 367
XX 229 NLLTTPFTVAMDWPKRPDAVCTWKQVEDEMURABYGGSFRFSSDAISTFTFTNLNLTQY 288
XX
XX 368 --SKEDMKEFNTTYHCLNEEIKABIEKKA-KVNSTHSKYGDLYKFKTGDGLVLMQPLI 424
XX 289 SLRSRVLDG-----CIGRDAREALDRMFARKYNATHIKVQPOPYLATGGFLIAYQPLL 342
XX
XX 425 QNRLLDAKKNLNNETYGR-RSRROAESTTDPMMEM-TGNAGAGGYSSENSITVAQVOYAY 482
XX 343 SNTLAE----LYVREYMRQDKPRNATPLREAPSANASVERIKTTSSTIEFARLQFTY 398
XX
XX 483 DNLIRIRNNILEDLSKAWCREOHAALVMNELSKINPTSVMSIYNRPVSAKRIGDVISV 542
XX 399 NHQQRHVDMLGRITAVAWCLQNLHELTMLNEARKLNPNATASATVGRVRARMGDVNAV 458
XX
XX 543 SNCIIVDQTSVSLHKSURLLASDEKFSRPPVTFKFMNDSTIYKGLGVNNELLFTTY 602
XX 459 STCVPAVDNPVIVQNSMR-VSSRPGTCYSRPLVSFRYEDQGLIEGQGENNELRLTRDA 517
XX
XX 603 LETQENETEFYQAKTDMYIYKNYEHKTKVPLSSITLDTFIALNFTLLENVDFKVIELY 662
XX 518 LEPCTVGHRRYFIFGGGYVFEYAYSHQLSRADVTVTVDITLNTMLDHEFVPLEVY 577
XX
XX 663 TRDEKRLSNVFDIETMREYNYAQRVSGRLKDLLDLSTNRNQFVD--AFGSLMDDLGA 719
XX
XX Db 578 TRHEIKDSGLLDYTEVQRRNQLHDLRPADI-DIVIRADANAAMFAGLCAFEFGMDLGRA 636
XX Qy 720 VGOVTNAVSGVATLFFSSIVTGFINKNPFQGMMLIIVVIGLVFAIYFLTKTKIYETA 779
XX Db 637 VGKVMGVGVGVV----SAVSGVSSFWNSNPFGLAVGLLVLAGLVAAFFAFRYVLQJQRN 692
XX Qy 780 PIKMIYPEIDKLEREGKSEIAPISBE-----ELERIVLAMHIHQNSHMET 826
XX Db 693 PMKALYPLTTTKBELKTSDPGCVGGEGBEGAGGGGFDEAKJAEAREMIRYMXLVSAMERTEH 752
XX Qy 827 KTRKDPKDSIL-TRAQNM-LRKR--SGYSNLKNAE 857
XX Db 753 KARKKGTSSALLSSKVTNMVLRKKNKARYSPHNED 787
XX
XX RESULT 32
XX AAP80914
XX ID AAP80914 standard; Protein; 904 AA.
XX AC AAP80914;
XX DT 25-MAR-2003 (updated)
XX DT 09-MAR-1992 (first entry)
XX DE Sequence of Herpes simplex virus (HSV) glycoprotein B (gB) 1.
XX KW Vaccine; herpes simplex virus; therapy.
XX OS Herpes simplex virus.
XX PN WO8802634-A.
XX PD 21-APR-1988.
XX PF 20-OCT-1987; 87WO-US02709.
XX PR 20-JUL-1987; 87US-0079605.
XX PR 20-OCT-1986; 86US-0921213.
XX (CHIR ) CHIRON CORP.
XX PI Burke LR, Pachl C, Valenzuela P;
XX WPI; 1988-119368/17.
XX N-PSDB; AAN80907.
XX Vaccine for treatment of herpes simplex virus - contains
XX recombinant HSV glycoproteins B and D
XX Example; Fig 4; 71pp; English.
XX
XX Prep. of recombinant gB and gD is described in WO88504587. The
XX amino acid sequence and DNA sequence for gB 1 presented in AAP80914
XX and AAN80907 differ from that originally presented in Table 1 of
XX International Publication No. WO 85/04587. The DNA sequence in Table
XX 1 contains an error in that an additional nucleotide (G) is listed
XX at position 607 which resulted in a shift in reading frame relative
XX to AAN80907 from which this nucleotide has been deleted.
XX (Updated on 25-MAR-2003 to correct PA field.)
XX
XX Sequence 904 AA;
XX
XX Query Match 18.3%; Score 833; DB 9; Length 904;
XX Best Local Similarity 28.2%; Pred. No. 1.6e-52;
XX Matches 245; Conservative 164; Mismatches 368; Indels 92; Gaps 27;
XX
XX 33 TSPNTATWSTESPLTGHYGTDDSSHGERGNENRDSSEQKNYIGSPSTFPYRVCSASG 92
XX 76 TPEPD----ANATVAAGHATLR--AHLREIKVENADAQ-----FVCPPT 115
XX 93 VGVDFRFQDTHVCPDASD-MVHSEGLIYKQNIIPFMRVRYKRVKVVTTSTV----- 144

```


Db 344 FNLQIHLSCQKBEARAIINRIYTRYNSHVRTGDIQTYLARGGVVVFQPLLSLSLA 403
 QY 430 -----DAKNKNNETYSRRSRQAESTTDPMMWETGAGCEYSSENSITVAQVQYADN 484
 Db 404 RLYLQELVRENTNHPKQHPTRNTRSRSSVPELRANT-----ITTSVEFAMLQFTYDH 460
 QY 485 LRIRINILEDLSKAWCREQHRALVWNLSEKINPTSVMSYINRPSAKRIGDVISVN 544
 Db 461 IQEHVNEMLARISSWCQQLNRERALSGLFPINPSALASTILDQKVARKILGDVISVN 520
 QY 545 CIVV-DQTSVSLHSLRLLSASDEKCSRPPVPTKFWNDSTIYKGLGVNNEILLTYYL 603
 Db 521 CPELGSDTRIILQNSMR-VSGSTTRCYSRPLISIVSLNGSGTVREGQLGTDNELIMSDLL 579
 QY 604 ETCQENTEYFQAKTDMVIYKNYEHKLTVPLESSITTLDTFALNFTLLENVDFKVIELYT 663
 Db 580 EPCVANKRYFLFGHHVYEDYRVREIAVHDVGMISTYVDLNLTLKLOREFMPLQVYT 639
 QY 664 RDEKRLSNVEDIETMFREYNYAQRVSGLRKDL-LDLSTNRNQFVDAFGSLMDDDLGAVGQ 722
 Db 640 RDELRTGLLDYSEIQRRNQMHSLRFYDIDKVQYDSGT---AIMQMAQFFQGLGTAGQ 696
 QY 723 TVNNAVSGVATLPSIIVTGFNFINKPFGCMMLIIVIGULFAIYFLTKTKIYETAPIK 782
 Db 697 AVGHVILGATGALLSTVHGFTTFLSNPFGALAVGLLVLAGLVAAPFAYRYVYLKLTSPMK 756
 QY 783 MIYPEIDK-LKE-REGKSEIA-----PISE-----EELERIV 812
 Db 757 ALYPLTTKGLQLEGMDPFAEKNATDTPLEELGDSQNTSPSVNSGDPDPKPREAQEMI 816
 QY 813 LAMHIQONSHMETKTRKDPKXDSILTRAQ-----NNMLRKRSGYSNLK 854
 Db 817 KYMTLVSAABEQESKARKNKTSALLTSRLTGLALRNRGYSVR 861
 RESULT 34
 AAW34553
 ID AAW34553 standard; Protein; 904 AA.
 XX AC AAW34553;
 XX DT 26-FEB-1998 (first entry)
 XX DE Herpes simplex virus type 2 glycoprotein B.
 XX KW Herpes simplex virus; HSV; type I; type II; glycoprotein B; vaccine;
 KW membrane protein; virus-specific glycoprotein;
 KW transmembrane anchor region.
 XX OS Herpes simplex virus type 2.
 XX PN US5648079-A.
 XX PD 15-JUL-1997.
 XX PF 08-DEC-1994; 94US-0351875.
 XX PR 20-OCT-1986; 86US-0921730.
 PR 06-APR-1984; 84US-0537784.
 PR 20-SEP-1990; 90US-0587179.
 PR 21-DEC-1992; 92US-0993415.
 PR 18-OCT-1993; 93US-0138717.
 XX PA (CHIR) CHIRON CORP.
 XX PI Burke RL, Pacht C, Valenzuela PDT;
 XX DR WPI; 1997-372022/34.
 DR N-PSDB; AAT93651.
 XX PT vaccines against herpes simplex virus infection - containing
 PT recombinant HSV glycoprotein B

PS Disclosure; Fig 4A-F; 33pp; English.
 XX The present sequence is from the Herpes simplex virus (HSV) type 2
 CC strain 333 and represents a glycoprotein B. HSV is a double stranded
 CC virus packaged within an icosahedral nucleocapsid enveloped within a
 CC membrane. The membrane includes a number of virus-specific
 CC glycoproteins, with glycoprotein B being one of the most abundant.
 CC Glycoprotein B from both HSV type I and type II are cross reactive. New
 CC vaccines comprising recombinantly produced glycosylated glycoprotein B
 CC that has a deletion of at least part of the transmembrane anchor region,
 CC in combination with a carrier and an adjuvant have been produced. The
 CC vaccines are used for immunising humans against HSV (HSV type 1 or
 CC HSV type 2) by vaccination before or after primary infection with HSV.
 XX Sequence 904 AA;
 SQ
 Query Match 18.2%; Score 829.5; DB 18; Length 904;
 Best Local Similarity 27.4%; Pred. No. 3e-52;
 Matches 243; Conservative 180; Mismatches 363; Indels 101; Gaps 28;
 QY 24 SLSIAETGVTSPTNTATWSTESP--LTGHYGTDDSHGE-RGNENRPSDEQKNKIYGSP 80
 Db 58 ALGAAGTDPKPKKKNKPKNPTPPAPAGDNATVAAGHATLREHLRDKAENTDANFY--- 114
 QY 81 STEPPYRVCSASGVGDVRFOTDHCPCDASD-MVHSGILLIYKONIIPFMRVRYKRYKV 139
 Db 115 -----VCPPTGATVVQEPQRCPTREGQNYTGIAVVFKNATPYKFKATMYKDV 168
 QY 140 TTSTV-----YNGIYSDSIQNTQHTFYKSIPEWE--TEKMDTIYQCFNSLRNLNTGNNL 189
 Db 169 TVSQVWFGHRYSQFMGIFED-----RAPVPEEVIDKINAKGVCRSTAKYVVRNLE 219
 QY 190 LTVVDREDDINMTVFLQPDGV-----TPDVKRYGSPQELYLEPGWFGVSGVRRRTTVN 241
 Db 220 TTAHFRRDDHETDMELKPNANAARTSRGWHITDLKYNPSRVEAF-----HRYGTTVN 270
 QY 242 CELMDMFARSNPDPDFVTATGTVEMSPFWS-GEDDH-ENKMKHKEPWFVSVINNY-KVY 298
 Db 271 CIVEEVDARSVPYDEFLVATGDFVYMSPFYGVREGSHTSHTSYAADRKFQVDGFFVARDL 330
 QY 299 DYQNRGTVPGLKTRIFLDREBYTSLWEKHLKNMSYCPCLTLWKAFYNGIQTQEHSGSVHFA 358
 Db 331 TTARATAP--TTRNLLTTPKFTVAMDVWPKRPSVCTMTKWQEVDEMLRAEYGGSPFSS 388
 QY 359 NDITASFTT-----SKEDMKEFNTTYHCLNEEIKAEIEKKA-KVNSTHSKYVDLKYF 410
 Db 389 DAISTFTTNLTTEYPLSRVDLGD-----CIGKQARDAMDRIFFARRYNATHIIVGQPOQY 442
 QY 411 KTDGGLYLVWQPLIQNRLLDAKNLNNETYSRRSRQAESTTDPMMWETGAGGEYSSE 470
 Db 443 LANGGLIAYQPLLSNTLAE-----LVVREHLREQSKPPTPTPPPGASANASVERIKTT 498
 QY 471 NSITVAQVQYADNLRIRINILEDLSKAWCREQHRALVWNLSEKINPTSVMSYINRPS 530
 Db 499 SSIEFARLQFTYHNIQRHVNDMLGRVAIAWCELQNHLELTWTEARKLNPANIASATVGR 558
 QY 531 VSAKRIGDVISVNCIIVDQTSVSLHSLRLLSASDEKCSRPPVPTKFWNDSTIYKGL 590
 Db 559 VSARMILGDVNAVSTVCPVAADNVIQNSMR-ISSREGACYSRPLVSFRYEDQOGLVEGQL 617
 QY 591 GVNNEILLTTLTLETQCENTEYFQAKTDMVIYKNYEHKLTVPLESSITTLDTFALNFTL 650
 Db 618 GENNELRLTRDALEPCTVGHRRYFTFGGGVYVFEESAYSQHSQLSRAIDITVTFIDLNITM 677
 QY 651 LENVDFKVIELYTRDEKRLSNVFDIETMFREYNYAQRVSGLRKDLDLSTNRNQF--VD 708
 Db 678 LEDHEFVPLEVYTRHEIKDSGLLDYTEVQRRNQLHDLRFADI-DTVIHADANAAMFAGLG 736
 QY 709 AFGSLMDDLG-AVGQTVNNAVSGVATLPSIIVTGFNFINKPFGCMMLIIVIGULFAIY 767
 Db 737 AFPEGMDLGRVAVGVVMIGVGVW-----SAVSGVSSVSMSPFPFALAVGLLVLAGLAAAF 792
 QY 768 FLTKTKIYETAPIKMIYPEIDK-LK-----EREGKSEIAPISE-EELERIVLA 814

```

Db      793 FAFYVMRLQSNPKALYPLTKELKNTNPDASGEGGDFDEAKLAAREMIRYMA 852
QY      815 MHIHQONSHMETKTRKDPKDSLT-RAQNML---RKRSGYNLKNAE 857
Db      853 VSAMERTHE---KAKKKGTSALLSAKVTDIMVMKRKRNTNYTQVFNKD 896

RESULT 35
AAP80915
ID AAP80915 standard; Protein; 905 AA.
XX
AC AAP80915;
XX
DT 25-MAR-2003 (updated)
DT 09-MAR-1992 (first entry)
XX
DE Sequence of Herpes simplex virus (HSV) glycoprotein B (gB) 2.
XX
KW Vaccine; herpes simplex virus; therapy.
XX
OS Herpes simplex virus.
XX
XX WO8802634-A.
PN
PD 21-APR-1988.
XX
PF 20-OCT-1987; 87WO-US02709.
XX
PR 20-JUL-1987; 87US-0079605.
PR 20-OCT-1986; 86US-0921213.
XX
PA (CHIR ) CHIRON CORP.
XX
PI Burke LR, Pacht C, Valenzuela P;
XX
DR WPI; 1988-119368/17.
DR N-PSDB; AAN80908.
XX
XX Vaccine for treatment of herpes simplex virus - contains
PT recombinant HSV glyco:proteins B and D
XX
PS Example; Fig 4; 71pp; English.
XX
CC Prepn. of recombinant gB and gD is described in WO88504587. The
CC amino acid sequence and DNA sequence for gB 1 presented in AAP80914
CC and AAN80907 differ from that originally presented in Table 1 of
CC International Publication No. WO 85/04587. The DNA sequence in Table
CC 1 contains an error in that an additional nucleotide (G) is listed
CC at position 607 which resulted in a shift in reading frame relative
CC to AAN80907 from which this nucleotide has been deleted.
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 905 AA;

Query Match
Best Local Similarity 27.3%; Pred. No. 8.2e-52; Length 905;
Matches 240; Conservative 176; Mismatches 338; Indels 125; Gaps 27.

QY 34 SPNTATWSTESPLTGHYTHDSHGE-RGNENRDSERQNIYGSPTFPYRCSAG 92
Db 78 TTPRPA-----GDNATVAAGHATLREHLRDIKAENTDANFY-----VCPPT 119
QY 93 VGDVFRPQDHYVCPDASD-MVHSEGLILYKQNIIPFMRVKRYKVTSTV----- 144
Db 120 GATVQVFEQPRRCPTREPCQNTGIAVFKENIAPYKFKATMYKDVTVSQVWFCHRY 179
QY 145 -YNGIYSDITNQHIFYSIEPWE--TEKWDIYQCFNSLRNLNTGNLLTYVDRDINMT 201
Db 180 QMGIFED-----RAPVFEVIDKINAKGVCSTAKYVRNNLETAFHRDDHETD 230
QY 202 VFLQPVQGV-----TPDKRYGSOPELYLEPGWFGSYRRRTTVCNCELMDFARSNP 253

```

```

Db      231 MELKPANAATRTSRGHTTDLKYNPSRVEAF-----HRYGTTVNCIVEEDARSVY 281
QY      254 PFDEFVATATGDTVMSPFWS-GEDDH-ENKMKHKEPWFVSVINNY-KVVDYQNRCTVPLGK 310
Db      282 PYDEFVLATGDFVYSSPFYGYREGSHTHTSYAADRFQVDGDFYARDLTKARATAP--T 339
QY      311 TRIELDREEXTLSWEKHLKNMSYCPLTLWKAIFYNGIQTEHSGSYHFFVANDITASFTT--- 367
Db      340 TRNLLTTPKFTVAWDVWPKRPSVCTMTKWQEVDEMLRSEYSGSPFRSSDAISTTFTTNLT 399
QY      368 ----SKEDMKFNTTYHCLNEEIKAEIEKKYA-KVNSTHSKYGLDKFKFDGGLYLWQOP 422
Db      400 EYPLSRVLDLGD-----CIGKDARDAMDRIFFARRYNATHIKVGPQFYLANGSLIAYQP 453
QY      423 LIONRLLDANKLNNETYSRRSRROASTTDPME-MTNGAGGAGSEYSENITVAQOYA 481
Db      454 LLSNTLAELVYR---EHLRQSRKPRNATAPLRGASANASVERIKTTSIEFARLOFT 510
QY      482 YDNLRIIRINILEDLSKAWCREQHRALVMNELSKINPTSVMSMIYRNPVSASAKRIGDVIS 541
Db      511 YNHIQRHVNDMLGRVAIAWCELQNHLELTWNARKLNPNATASATVGRRRVSARMLADVMA 570
QY      542 VSNCIIVVDQTSVSLHKSRLLSASDEKCFSPRPVTPFKMNDSTIYKQGLVNNELLITTT 601
Db      571 VSTCPVAADNVIQNSMR-ISSRPGACYSRPLVSFRYEDQGPLVEGQSGENNELRLTRD 629
QY      602 YLETCQENTEVYFOAKTDMYIKNYEHLKTVPLSSITTLDTFIALNFTLLENVDFKVEL 661
Db      630 AIEPCTVGHRRYFTFGGYVYFESAYSHQUSRADITTVSTFDLNTMLDHEFPVLEV 689
QY      662 YTRDEKLSNVFDIETMFRYNYAQRVSGRLKDLDDLSTNRNQF--VDAFSGSLMDDLG- 718
Db      690 YTRHEIKDSGLLDYTEVQRNQLDLRFADI-DTVIHADANAAMFAGLGAFFEGMGDLGR 748
QY      719 AVGTVMNAVSGVATLFSISIVTGFINFKNPFQGMILIIIVIGVLFAIYFTKTKIYET 778
Db      749 AVGVWVGIVGGVV-----SAVSGVSSFMSPNPFALAVGLLVLAGLAAFAFFAFRYMRLQS 804
QY      779 APIKWIYPEIDKLKEREKSEIAPISBEELERIVLAMIHQONSHMETKTRKDPKDSILT 838
Db      805 NPMKALYP---LTTKELKNPTNPDASGEG-----EGGHFD-----EAKLA 842
QY      839 RAQNMLR-----KRSGYSNLKNASVEML 862
Db      843 EAREMIRYMALVSAMERTEHKAKKGTSAULSAKVTDVW 881

RESULT 36
AAP79459
ID AAP79459 standard; Protein; 873 AA.
XX
AC AAP79459;
XX
DT 25-MAR-2003 (updated)
DT 16-JAN-1996 (first entry)
XX
DE Infectious Laryngotracheitis virus gB glycoprotein.
XX
KW Infectious Laryngotracheitis virus; gB; glycoprotein; antibody;
KW vaccine; immunoassay; detection; identification; avipox; chicken;
KW fowl; probe; antisense; gene expression.
XX
OS Gallid Herpesvirus 1.
XX
XX Key Location/Qualifiers
XX Key Location 656..659
XX Misc-difference /note="These amino acids are missing from the
XX FT decoded polypeptide sequence given in the
XX FT specification."
XX FT
XX PN US5443831-A.
XX PD 22-AUG-1995.

```

XX 19-NOV-1993; 93US-0156866.
 XX 29-OCT-1991; 91US-0788123.
 PR 19-NOV-1993; 93US-0156866.
 XX (UYDE) UNIV DELAWARE.
 XX Keeler CL, Poulsen DJ;
 FI WPI; 1995-302091/39.
 XX N-PSDB; RAO97351.
 DR Isolated Infectious Laryngotracheitis Virus GB gene - used to
 PT develop prods. for detection and for vaccine(s), partic. for
 PT immunising fowl.
 XX
 PS Claim 1; Columns 11-16; lipp; English.
 XX
 CC The infectious laryngotracheitis virus (ILV) GB glycoprotein
 CC nucleic acids can be used as probes, for antisense control of GB
 CC gene expression or for production of GB polypeptides. The GB
 CC polypeptides can be used for the production of antibodies (both
 CC used in immunoassays) and in vaccines. Recombinant avipox virus
 CC expressing the ILV GB can be used for immunising fowl, especially
 CC chickens.
 CC (Updated on 25-MAR-2003 to correct PF field.)
 XX
 SQ Sequence 873 AA;
 Query Match 18.1%; Score 821; DB 16; Length 873;
 Best Local Similarity 27.6%; Pred. No. 1.2e-51;
 Matches 220; Conservative 158; Mismatches 310; Indels 108; Gaps 23;
 QY 57 SHGERGNNEPD--SEEQNKIY-----QSPSTPYR---VCSASGVGDVRFQTDHV 104
 DB 24 SHGIAGIDPDRTASMDVGKISFSEATGSGAPKEPQIRNRFACSSPTGASVARLAQPRH 83
 QY 105 CPDASDMVH-SEGILLIYQKQIIPFMRVRKYRKVWTTSTVYNGIYSDSITNQHTFYKSI 163
 DB 84 CHRHADSTNTEGIAVFKNIAPYVENVTLIYKHITVTWALFSPQITNEVTRVPI 143
 QY 164 EPWETKMDTYIQCFNSLRLNTGNGLITYVDRDINMTVFLQ-----VDGVTPDV 214
 DB 144 DYHEIVRIDRSGECSKATYHKNFMEFAYDNDEREKKLPLVPSLLRSTVSKAPHTTNT 203
 QY 215 KYGSGQELYLEPCWFGWYSRRRTVNCELMDMFARSNPDPFVTATGDTVMSFWSG 274
 DB 204 KRHOTL-----GYRTSTSDVCVEYLQARSVPYDFGMATGDTVEISPFYT- 250
 QY 275 EDDHENKMHKPFVSVINNYK---VVVDYQNR---GTVPLGKTRIFLDREYVTLSEKH 327
 DB 251 -----KNTGPRRHSVYRDYRFLIEANYQVRDLETQIRPPKENFLTDEQFTIGHAM 304
 QY 328 LKNMSYCLPLWKAIFYNGIQTEHSGSVHFVANDITASTTSKEDMKFNNT---YHCL- 382
 DB 305 EEKESVCTLSKWIEVPAVRVSKNSYHFSKDMTMTFFSGKQ---PFNISRHLAECPV 361
 QY 383 ---NEEIKAELEKKAQVNSHSHKYGDLKYFKTDGGLYLVNQPILQNRLLD---AKNKLN 436
 DB 362 TIATEADIGTFARKY---SSTHVRSQDIEIYVLGSGFLIAFQKLMHGLAEWYLEEQ 418
 QY 437 NETYSRRSRQAESTTDPMMEMTNGAGGEYSSENSITVAQVQVAYDNLAIRINNILEDL 496
 DB 419 NHLPRGRERQAAGRRTASLQ--SGPQDRLTHISSATFAMLQFAYDKIQAHVNNELIGNL 476
 QY 497 SKAWCRQHRALVWNELSKINPTSVMSIYNRPVSAKRGDVISVNSNCIIVDQTSVSLH 556
 DB 477 LEAWCELQNRQLIIVWHEMKLNPNSLMTSLFGQVPSARLLGDIVAVSKCIEPIENIRMQ 536
 QY 557 KSLRLLSASDEKCFSPRPVTFKM-----NDSTIYKQGLGVNNEILLTITYL 603
 DB 537 QSMR-MFGDPTCYTRPVLIFRYSSESPESQFANSSTENHNDILQGLGEHNEILQGNRLI 595

QY 604 ETCQENTYFQAKTDMVIYKNYEHKLTVPSSLSTITLDTIALNFTLLENVDFKVIELYT 663
 DB 596 EPCWINHRRYFLLGENVLLYEDYTFVRQVNASIEEVSIFINLNATILEDLDFVPVEVYT 655
 QY 664 RDEKRLSNVEDIETMPREYNYYAQR-----VSGURKDLDDLSTNRNQFVDAFGSLMDD 716
 DB 656 REELRDTGLNYDDVVRYQIYNKRPRDIDTVIRGDRDAI-----FRAIDAFGNTLGE 710
 QY 717 LG-AGQTVNAVSGVATLFFSSIVTGFINKNPFQGM-----LMIIVIGVLPAIYFLT 770
 DB 711 VGRALGTVVMTAAAVI-----STVSGIASFLSFAALGIGIAVVSIIIGLAFKYVMN 766
 QY 771 KTKIYETAPIKMIYP 786
 DB 767 LKSN-----PVQVLP 777
 RESULT 37
 AAP70769
 ID AAP70769 standard; protein; 973 AA.
 XX
 AC AAP70769;
 XX 25-MAR-2003 (updated)
 DT 11-MAR-1991 (first entry)
 XX
 DE Glycoprotein B of herpes simplex virus type 1.
 XX
 KW Herpes simplex virus; glycoprotein B;
 XX
 OS Herpes simplex virus.
 PH Key Location/Qualifiers
 FT Region 11..26
 FT /label=hydrophobic putative signal
 FT Region 86..88
 FT /label=potential glycosylation site
 FT Region 140..142
 FT /label=potential glycosylation site
 FT Region 429..431
 FT /label=potential glycosylation site
 FT Region 558..560
 FT /label=potential glycosylation site
 FT Region 803..840
 FT /label=anchor region
 FT Region 844..864
 FT /label=anchor region
 FT Region 888..890
 FT /label=potential glycosylation site
 FT Region 957..959
 FT /label=potential glycosylation site
 XX EP236145-A.
 PN 09-SEP-1987.
 PD 09-MAR-1987; 87EP-0302001.
 PF 07-MAR-1986; 86GB-0005646.
 PR 01-SEP-1986; 86GB-0021081.
 PR 16-DEC-1986; 86GB-0029988.
 PR 09-MAR-1987; 87GB-0024676.
 XX
 PA (COGE-) COGENT LTD.
 PA (SMIT/) SMITH G L.
 PI Smith GL, Barrell BG, Cranage MP;
 XX WPI; 1987-251648/36.
 DR
 XX Prodn. of antigenic, human cytomegalovirus polypeptide(s) - by
 PT expressing recombinant vector, able to induce neutralising antibodies,
 PT

Db	122	RAPHQNTGEGIAVIFKENIADPKFATMYXKDVTSQVWFGRHSYQFNGIFED-----	175
Qy	158	TFYKSGIEPWE--TEKMDTIYQCFNSLRNLNTGCGNLLTYVDRDDINMTVFLOPVDGVTPDPK	215
Db	176	---RAPVPFEEVIDKINARGVCRSTAKYVRNNWESTAFHRDDDESDMKLKPAKAAAT-RTS	231
Qy	216	RYGSOPELYLEFGFWGYSRRRTTVNCELMDFARSNPDPDFVTATGDTVMSPFWSGE	275
Db	232	RGWHTTDLKYNPSRIEAFHRYGTTVNCIVEEVEARSVPYDFEVLATGDFVYMSFFYGYR	291
Qy	276	D--DHENKMKHEKPMFWVSVINNY-KVVDQNRQRTVPLGKTRIFLDRREYTLSEKHLKNMS	332
Db	292	DGAHAHTAYAADRPRQVDGYVERDLSTGRRASTP--ATRNLLTTPKFTVGTGMDWAPKRPS	349
Qy	333	YCPILTLKFAFYNGIQTEHSGSVHFVANDITASFTTTSKEDMKFNFTY-----HCLNE	384
Db	350	VCTLTFKWQVEMLRAEYGPFRFSSSALLSTFTTNR-----TEYALSRVDLGDGCVGR	402
Qy	385	EIKAEIEKKY-AKVNSTHSGYKGLKFKYKFTDGGLYLVWQPLIQRLLDAKNKLNNETYSRR	443
Db	403	EAREAVDRIFLRRYNGTHVKVQVYLLATAGFLAYQPLLSNGLV-----ELYVRE	454
Qy	444	SRROAE-----STTDPMMEMTNGAGGYSSENSITVAQOVAYYNLRINNN	491
Db	455	LLREQEGRPGDAATPKPSADPPDVE-----RIKTTSSVEFARLQFTYDHIQRHVD	506
Qy	492	ILEDLSKAWCREOHRALVMEKLSINPTSVMSIYNRPVSAKRIGDVISUSNCIIVDOT	551
Db	507	MLGRITAIWCELQNHLETLWNEARKLNPAITASATVGRVRSA RMIGDVMNAVSTCPEVTPD	566
Qy	552	SVSLHKSRLLSASDEKCFSPPTVFKFMNDSTIYKGQLGVNNEILLTITYLETQENTE	611
Db	567	NVIQNSMR-VPARPCTCYSRPLVSFRYEEGGLVEGQLGEDNEIRLERDALEPCTVGHR	625
Qy	612	YYQAKTDMYIYKNYEHUKTVPLSSITLDTFIALNFTLLENVDPFKVIELYTRDEKRLSN	671
Db	626	RYFTGAGYVYPEDYAVSHQLGRADVTTVSTFINLNLTWLEDHFVPLEVYVTRQEIKDSG	685
Qy	672	VFDIETWPERYNYAQRVSGLRKOLLDLSTNRNQFVDAFGSLMDDLGAQGTQVNVASGV	731
Db	686	LLDYTEVQRRNQLHALRFADI-DTVIKADAHAPLFAGLY-SFPEGLGDVGRVAVGVNMG	743
Qy	732	ATLFSSIVTGFINFKNPFGGMLMIIVIGVLFAIVFLTKTKIYETAPIKMIYPEIDKL	791
Db	744	VGCVVSAVSGVSSFLSNPFGALAVGLLVLAGLAAFAFRYMRLQRPMPKALYPLTTKE	803
Qy	792	KEREKGE-----IAPISEEELERIVLA-----MHIHQNSHMETKTRKDPKDSIL	837
Db	804	LKSDGSRGDCGDGASGGGEEDFEAKLAQAREMIRYALVSAMERTEHKARKKTSALL	863
Qy	838	T-RAQNM-LRKRS--GYSNLKNARESVM	861
Db	864	SAKVTNMVMRKRAKPRYSPLGDTDEEL	891
RESULT 39			
AAW70293			
ID	AAW70293 standard; Protein; 891 AA.		
XX	AC		
XX	AAW70293;		
DT	06-NOV-1998 (first entry)		
XX	Simian herpesvirus B gB glycoprotein sequence (UL27) .		
DE	Simian herpesvirus B gB glycoprotein; UL27; ICP protein; UL28;		
XX	differential diagnostic test; immunoassay; antibody.		
KW	Simian herpesvirus B.		
XX	US5767265-A.		
XX			
FN			
XX			

Qy 204 LQPV DGVTPDV KRYGSQPELYLEPGWFWGSYRRRTTVNCELMDMFARSNPFFDFVTATG 263

```

Db 244 LRPSRLNALGTRAWHTTNDYTKIG-AAGFYQTGTSVNCIVEEARSVPYDPSALSTG 302
Qy 264 DTVEKSPFWS-GEDDHKNMHEKPFVSVINNYKVVDYQNRGTVPGLKTRIFELDREYTL 322
Db 303 DIVYMSPPFGLREGAGEQIGVAPGRFQVYHYPIIDLSRLRASESVTRNLFRTHTFV 362
Qy 323 SWEKHLKNMSYCLTLWKAIFYNGIQTE-HSGSYHFVANDITASFTT--SKEDMKEFNTTY 379
Db 363 AMDWAPKTRRVCSLAKWREAEEMTDETRDGSFRFTSRALGASFDVDTQLDLQRVHLG- 421
Qy 380 HCLNEIKAEIEKKY-AKNVSTHSGYDGL-KYFKTDGGLYLVWQPLIONRLDA-KNKLN 436
Db 422 DCVLREASEAIDAIYRRRYNSTHVLADRPEVYLARGGFVAFRPLISNELAQLVARELE 481
Qy 437 N-----ETVSRSSRQ--AESTTDPMMETGNGAGGEYSSENSITVAQVQYAYD 483
Db 482 RLGLAGVCPAPAAARRARRPGFAGTPEP---PAVNGTGHRLRTTGSAAEFARLOFTVD 538
Qy 484 NLRIIRINILEDLSKAWCREQHRALAVWNELSKINPTSVMSMIYNRPVSAKRIGDVISV 543
Db 539 HQAHVNDMLGRIAAAWCLOKNDRTLNSEMSRLNPSAVATAALQORVCARMLGDVMAIS 598
Qy 544 NCIVVDQTSVSLHLSLSDSEKCFSPRPVTFKPMNDSTIYKQGLGVNNEILLTLYL 603
Db 599 RCVEV-RGGVYVQNSMR-VFGERGTCYSRPLVTFEH-NGTGVIEGQLGDDNELLISRDLI 655
Qy 604 ETCQENTYFYFOAKTDMYIKNYEHLKTVPLSSITTLDTFFIALNFTLLENVDVKVIELYT 663
Db 656 EPTGNHRRYFKLGSGYVYEDYNYVRMVEFPE--TISTRVTNLTLNLTLEDREFFLEVYT 713
Qy 664 RDEKLSNVFDIETMFREYNYA-----QRVSLGRKLLDLSTNRNOFVDAFGSLMDL 717
Db 714 REELADTGLDYSEIQRRNQLHALKFYDIDRVVVDHNVLLRGIANEF-----OGL 765
Qy 718 GAVGQTVNAGSVATLSSIVTGFINFKNPFGMLMIIVIGVLFYFAIYFLTKTKIYE 777
Db 766 GDVGAAGVKWLGCATCAVISAGVGMVSLNSPFGALGALLVLGLVAAFLAYRHSRLR 825
Qy 778 TAPIKWIPEIDKLEREGKSEIAPISEELER-----IVLAMIHQONSHWEKTK 828
Db 826 RNPKALIPVTTKTUKEDGVDE-GVDDEAKLDQARDMIRYMSIVSAL---EQQEHKARKK 881
Qy 829 RKDPKDSILTRAQNMLRKRSGYSNLKNAE 857
Db 882 NSGPA-LLASRVGAMATRRHYQRIESED 909

RESULT 45
AAW36051
AC AAW36051 standard; Protein; 1086 AA.
XX
XX AAW36051;
XX
DT 15-JUL-1998 (first entry)
XX
DE Hybrid Marek's disease virus/M. gallisepticum 40 K-C protein.
XX
KW Chimeric; Marek's disease virus; outer membrane protein; fusion protein;
XX
OS Chimeric - Marek's disease gammaherpesvirus.
XX
OS Chimeric - Mycoplasma gallisepticum.
FH Key Location/Qualifiers
FT 1..672
FT /note= "derived from Marek's disease virus gB protein"
FT 693..1086
FT /note= "derived from M. gallisepticum antigen"
XX
PN W09736924-A1.
XX
PD 09-OCT-1997.

```

```

XX 28-MAR-1997; 97WO-JP01084.
XX
XX 29-MAR-1996; 96JP-0103548.
XX
XX (JAPG) NIPPON ZEON KK.
XX
XX Saito S, Tsuzaki Y, Yanagida N;
XX
XX WPI: 1997-503046/46.
XX
XX N-PSDB; AAT96596.
XX
XX Fusion protein comprising herpes virus outer membrane protein and
XX antigenic polypeptide - for prevention of infection by Mycoplasma
XX gallisepticum, especially in poultry
XX
XX Disclosure; Page 22-30; 51pp; Japanese.
XX
XX This sequence represents the chimeric protein 40 K-C which comprises a
XX fragment of the Marek's disease virus outer membrane protein gB fused
XX to an antigenic protein from Mycoplasma gallisepticum. The chimeric
XX protein can be used in recombinant live vaccines for prevention of
XX infection by Mycoplasma gallisepticum, especially as the outer membrane
XX protein shows antigenicity in poultry.
XX
XX Sequence 1086 AA;

Query Match 16.6%; Score 757; DB 18; Length 1086;
Best Local Similarity 25.9%; Pred. No. 8.7e-47;
Matches 225; Conservative 157; Mismatches 374; Indels 114; Gaps 22;

Qy 51 YGTHDSHGERGNENRDSBEQNKNIYGSPTFPVVCSSAGVGVDVFRFQTHVCPDASD 110
Db 17 YGTNSP-----STQNTSVRSVSVLSSEESTFYLCPPVGVSTVIRLEPRKCPERK 71
Qy 111 MVH-SEGILLIYKQNIIPFMRVRKRVKVTSTVYNGIYSDSITNQHTFYKSIPEWE-T 168
Db 72 ATENGEGAILLFENISPYKFKVLYKNIITQTTWTGTYRTNRYTDRTPVSEIT 131
Qy 169 EKMDTIYQCFNSRLNTGGNLLTVVDRDDINMTVLPQVGVTPDKRYGSGOPELYLSPG 228
Db 132 DLIDGKGRCSKARYLRNNVYVEAFDRDAGEKQVLLKPKFNTPESAWHTTNETYT--- 188
Qy 229 WFWGS---YRRTTVNCELMDMFARSNPDPDFVTATGDTVEMSPFW--SGEDDHENKWH 283
Db 189 -WVGSPWYRTGTSGVCIIVEEMDARSVPFYSFAMANGDIANISFFYGLSPPEAAEPMG 247
Qy 284 EKPWFVSVINNYKVVDYQNRGTVPGLKTRIFLDREYTLSEKHLKNMSYCLTLWKAFY 343
Db 248 YPQDNFKQLDSYFSMDLDKRRKASLPVKRNFLLTSHFTVGVWDWAPKTRVCSMTKKKEVT 307
Qy 344 NGIOTHSYHGFVANDITASFTTSKEDMKEFNTTY-----HCLNEIKAEIEKKY 394
Db 308 EMLRATVNGRYRFARELSATFIS-----NTTEFPNRIILCQCIKREAAAEIQIF 359
Qy 395 -AKVNSTHSGYDGLKYFKTDGGLYLVWQPLIQ-----NELLDAKNKLN 436
Db 360 RTKYNDSHVKVGHVQYFLALGGFIVAYQPVLSKLAHMYLRELNRDRTDMLDLVNNKH 419
Qy 437 ----NNTYSSRRSRQAEESTTDPMMETGNGAGGEYSSENSITVAQVQYADNLRIRIN 491
Db 420 AIYKKNATLSRLRDIRNAPNRKITLDDTTA---IKSTSSVQFAMQLFYDHIQTHIND 476
Qy 492 ILEDLSKAWCREQHRALAVWNELSKINPTSVMSMIYNRPVSAKRIGDVISVNCIVVDQT 551
Db 477 MFSRIATAWCELQNRLELVWHEGIKINPSATATLGRVAAKMLGDVAAVSSCTAIDAE 536
Qy 552 SVSLHLSLRLSLSADEKCFSPRPVTFKPMNDSTIYKQGLGVNNEILLTLYLFCQENTE 611
Db 537 SVTLQNSMRVIT-STNTCYSRPLVLFYSYGENQGNIQGLGENNELLPLEBAVEPCSANHR 595
Qy 612 YYFOAKTDMYIKNYEHLKTVPLSSITTLDTFFIALNFTLLENVDVKVIELYTRDEKLSN 671

```

Db 596 RYFLFGGYALFENYFVKWVDAADIQIASTFVELNLTLLQDREILPLSVVYTKRELRDVG 655
Qy 672 VFDIETMREYNYAQRVSLRKDLLDLSTNRNOFVDAFGSLMDDDLGAVGQTVVNAVSGV 731
Db 656 VFDYAEVARRNQLHELKFDINK-VIEVDYNY-----AGLOEFGCMSITKKDA---- 702
Qy 732 ATLFSSIVTGFINFIKNPFGG-----MLMIIWIGVLFAIYFLTKTKIYETAPIKM 783
Db 703 -----NPNNGQTOLEAARMELTDLINAKAMTTLASLODYAKI---EASLSS 744
Qy 784 IYPEIDKIKEREKGESEIAPISEEELERIVLAMHIHQONSHMETKTRKDP-----KD 834
Db 745 AYSEAEVNNLN-----ATLEQLKMAKTNLESAINQANTDKTTFDNEHPNLVEAYKAKLT 800
Qy 835 SILTRAQNMUKRS-GYSNLKNAESVEMLN 863
Db 801 TLEQRATNLEGLSTAYNQIRN-NLVDLYN 829

Search completed: October 28, 2003, 15:33:08
Job time : 99 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 15:31:05 ; Search time 81 Seconds
(without alignments)
1788.314 Million cell updates/sec

Title: US-10-055-364-24
Perfect score: 4547
Sequence: 1 MAGSLKRGSLVLAQWLYQV.....KRGYSYLNKNAESVEMLNTL 865

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 629382 segs, 167460630 residues

Total number of hits satisfying chosen parameters: 629382

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US05_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4547	100.0	865	14	US-10-055-364-24
2	2473	54.4	793	14	US-10-055-364-55
3	2438	53.6	830	14	US-10-055-364-45
4	1940	42.7	824	14	US-10-055-364-43
5	1918	42.2	823	14	US-10-055-364-37
6	1916	42.1	792	14	US-10-055-364-42
7	1902	41.8	808	14	US-10-055-364-38
8	1898.5	41.8	791	14	US-10-055-364-41
9	1871.5	41.2	818	14	US-10-055-364-44
10	1844.5	40.6	829	14	US-10-055-364-46
11	1780	39.1	831	14	US-10-055-364-39
12	1775	39.0	844	14	US-10-055-364-40
13	1128	24.8	907	15	US-10-223-538-2
14	1002	22.0	195	14	US-10-055-364-2
15	844	18.6	904	9	US-09-894-998-18

16	844	18.6	904	12	US-10-200-562-18
17	844	18.6	904	12	US-10-237-551-18
18	844	18.6	904	15	US-10-121-988-18
19	754	16.6	1086	9	US-09-147-052-4
20	442	9.7	359	14	US-10-055-364-48
21	350	7.7	206	14	US-10-127-733-2
22	127	2.8	526	9	US-09-815-242-13696
23	126.5	2.8	3899	15	US-10-171-311-4
24	126.5	2.8	3907	15	US-10-171-311-2
25	126.5	2.8	3917	15	US-10-171-311-8
26	126.5	2.8	3925	15	US-10-171-311-6
27	125	2.7	1009	9	US-09-815-242-12141
28	124	2.7	996	9	US-09-815-242-5251
29	124	2.7	1565	15	US-10-242-056-59
30	123	2.7	792	9	US-09-815-242-12327
31	120	2.6	748	9	US-09-815-242-12792
32	119	2.6	2175	11	US-09-978-244A-4
33	117.5	2.6	602	12	US-10-032-585-7379
34	117.5	2.6	1025	10	US-09-486-734A-4
35	117	2.6	1639	15	US-10-087-464-10
36	116.5	2.6	1169	9	US-09-815-242-13448
37	115.5	2.5	1167	9	US-09-815-242-11522
38	115	2.5	813	10	US-09-732-180-7
39	115	2.5	813	12	US-10-054-399A-29
40	115	2.5	844	10	US-09-732-180-2
41	115	2.5	844	12	US-10-054-399A-24
42	114.5	2.5	1169	9	US-09-815-242-13693
43	114.5	2.5	2297	15	US-10-245-802-20
44	114	2.5	2184	12	US-10-304-095-6
45	113.5	2.5	1844	15	US-10-242-056-53

ALIGNMENTS

RESULT 1
US-10-055-364-24
; Sequence 24, Application US/10055364
; Publication No. US20020155433A1
; GENERAL INFORMATION:
; APPLICANT: Patience, Clive
; TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use
; FILE REFERENCE: 61750-379
; CURRENT APPLICATION NUMBER: US/10/055.364
; CURRENT FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: US/09/612,204
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US/60/142,736
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: US/60/168,532
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 24
; LENGTH: 865
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Deduced amino acid sequence of porcine gamma herpesvirus gpB gene
US-10-055-364-24

Query Match	100.0%;	Score 4547;	DB 14;	Length 865;
Best Local Similarity	100.0%;	Pred. NO. 0;		
Matches 865;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MAGSLKRGSLVLAQWLYQVALYSLAETGVTSPPTATWTSTESPLTGHYTHDSHGE	60	
Db	1	MAGSLKRGSLVLAQWLYQVALYSLAETGVTSPPTATWTSTESPLTGHYTHDSHGE	60	
Qy	61	RGNNEKRSSEQNKNVYGVSPSTFPYRVCSAGSGVDVFRFQTDHVCPCDASDMVHSEGI	120	
Db	61	RGNNEKRSSEQNKNVYGVSPSTFPYRVCSAGSGVDVFRFQTDHVCPCDASDMVHSEGI	120	

```
Qy 121 YKQNIIPMRVRKYRVKVTSTVYNGIYSDSIITNOHTFYKSIIEPWETEKMDTIYOCFNS 180
Db 121 YKQNIIPMRVRKYRVKVTSTVYNGIYSDSIITNOHTFYKSIIEPWETEKMDTIYOCFNS 180
Qy 181 LRLTGNGLTYVDRDINMTVFLQPDVGVTPDKRYGSOPELYLEPGWFGSYRRRTTV 240
Db 181 LRLTGNGLTYVDRDINMTVFLQPDVGVTPDKRYGSOPELYLEPGWFGSYRRRTTV 240
Qy 241 NCELMDFARSNPPDFPVATGDTVEMSPFWGSDHDKMHEKMPFVSINNYKVVDY 300
Db 241 NCELMDFARSNPPDFPVATGDTVEMSPFWGSDHDKMHEKMPFVSINNYKVVDY 300
Qy 301 QNRGTVPGLKTRIFLDREEVYLSWEKHLKMSYCPLTLWKAFYNGIOTESHGSHVHFVAND 360
Db 301 QNRGTVPGLKTRIFLDREEVYLSWEKHLKMSYCPLTLWKAFYNGIOTESHGSHVHFVAND 360
Qy 361 ITASFTTSKEDMKEFNTTYHCLNEEIKAEIEKKYAKVNSTHSHKYGDLKYFKTDGGLYLW 420
Db 361 ITASFTTSKEDMKEFNTTYHCLNEEIKAEIEKKYAKVNSTHSHKYGDLKYFKTDGGLYLW 420
Qy 421 QPLIQNRLLDKAKNINNETYSRRRROAESTTDPMMEMTNGAGGEYSSENSITVAQVY 480
Db 421 QPLIQNRLLDKAKNINNETYSRRRROAESTTDPMMEMTNGAGGEYSSENSITVAQVY 480
Qy 481 AYDNLRIIRINNILEDLSKACREOHRAALVWNLKINPTSVMSMIYNNRPVSAKRIGDVI 540
Db 481 AYDNLRIIRINNILEDLSKACREOHRAALVWNLKINPTSVMSMIYNNRPVSAKRIGDVI 540
Qy 541 SVSNCIVVDOTSVSLHSLRLLSASDEKCFSRPPVTFKFWNDSTIYKQGLGVNNEILLTT 600
Db 541 SVSNCIVVDOTSVSLHSLRLLSASDEKCFSRPPVTFKFWNDSTIYKQGLGVNNEILLTT 600
Qy 601 TYLETQENTYFQAQTDYIYKNEYHLKTVPLSSITTLDTFIALNFTLLENVDFKVI 660
Db 601 TYLETQENTYFQAQTDYIYKNEYHLKTVPLSSITTLDTFIALNFTLLENVDFKVI 660
Qy 661 LYTRDEKRLSNVDFIETMFREYNYAQRVSGRLKDLDDLSTNRNQFVDAFGSLMDDLAV 720
Db 661 LYTRDEKRLSNVDFIETMFREYNYAQRVSGRLKDLDDLSTNRNQFVDAFGSLMDDLAV 720
Qy 721 GQTVVNAVSGVATLFSSIVTGFINFKNPGGMLMIIVJGVLFPAIYFLTKTKIYETAP 780
Db 721 GQTVVNAVSGVATLFSSIVTGFINFKNPGGMLMIIVJGVLFPAIYFLTKTKIYETAP 780
Qy 781 IKMIYPIIDKLKEREKSEIAPISEBELRIVLAMHIHQONSHMETKTRDKDPSILTRA 840
Db 781 IKMIYPIIDKLKEREKSEIAPISEBELRIVLAMHIHQONSHMETKTRDKDPSILTRA 840
Qy 841 QNMLRKESGYSLNKAESVEMLNTL 865
Db 841 QNMLRKESGYSLNKAESVEMLNTL 865

RESULT 2
US-10-055-364-55
; Sequence 55, Application US/10055364
; Publication No. US20020155433A1
; GENERAL INFORMATION:
; APPLICANT: Patience, Clive
; TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use
; FILE REFERENCE: 61750-379
; CURRENT APPLICATION NUMBER: US/10/055,364
; PRIOR FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: US/09/612,204
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US/60/142,736
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: US/60/168,532
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 55
; LENGTH: 793
```

```
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Portion of Acelaphine herpesvirus.
US-10-055-364-55

Query Match 54.4%; Score 2473; DB 14; Length 793;
Best Local Similarity 57.8%; Pred. No. 9.6e-203;
Matches 468; Conservative 130; Mismatches 169; Indels 46; Gaps 8;

Qy 74 KNIYSPSPFPYRVCASGVGVFRFQTDHVCVDASDMVHSEGLLIYKONIIPFMRVR 133
Db 1 KGHSDPSAFPFRVCASNIGDIFRQTSHCNTKDKHEHNEGLLIIFKENIYPPYFKVR 60
Qy 134 KYRKVVTSTVYNGIYSDSIITNOHTFYKSIIEPWETEKMDTIYOCFNSLRLNTGNLTYY 193
Db 61 KYRKIVTSTVYNGIYADAVTNOHVFSKSPVIETRRMDTIYQCYNSLDVTGNGNLVYT 120
Qy 194 DRDDINMTVFLQPDVGVTPDKRYGSOPELYLEPGWFGSYRRRTTVNCELMDFARSNP 253
Db 121 DNDGSNMTVDLQPDVGLSNSVRRYHSQPEIHAEPGLLGGYRRRTTVNCEVTETDARAVP 180
Qy 254 PFDFVTATGDTVEMSPFW-----SGEDDHENKMKHEKMPFVSINNYKVVDYONRG 304
Db 181 PFRYFITNIGDTIEMSPFWKAWNETEFGSGEDPRT-----LIVAKDYRVVDYKFRG 231
Qy 305 TVPLGKTRIFLDREEVYLSWEKHLKMSYCPLTLWKAFYNGIOTESHGSHVHFVANDITAS 364
Db 232 TPOGHTRIFVDKEEYTLSSWAQOFRNISYCRWAHWSFONAIKTEHGKSLHFVANDITAS 291
Qy 365 FITSKEDMKEFNTTYHCLNEEIKAEIEKKYAKVNSTHSHKYGDLKYFKTDGGLYLWQPLI 424
Db 292 FYTPNTQTRVGLGKHVCLANTIESELKSLRAKLVNDTHSPNGTAQYLYLTNGGLLWQPLV 351
Qy 425 QNRLDANKLN-----NETYSRRRROAESTTDPMMEMTNGAGGEYSSENSITV 475
Db 352 QOKLLDAKGLLDVAKKQONTTITTTTTRRRRQRRSVS-----SGIDDDVYTAESTILL 403
Qy 476 AQVQAYDMLRIRINNILEDLSKACREOHRAALVWNLKINPTSVMSMIYNNRPVSAKR 535
Db 404 TQIQFAYDTLRAQINNVLEBELSRACREOHRAASLMMNELSKINPTSVMSIYGRPVSAKR 463
Qy 536 IGDVISVSNCIIVVDOTSVSLHSLRLLSA-SDEKCFSRPPVTFKFWNDSTIYKQGLGVN 594
Db 464 IGDVISVCHVVDVQDSVSLHRSRMRVPGDKTHECYSRPPVTFKFINDSHLYKQGLGVN 523
Qy 595 EILLTTTLETQENTYFQAQTDYIYKNEYHLKTVPLSSITTLDTFIALNFTLLENV 654
Db 524 EILLTTTAVEICHENTEHYFQGNMNYFYKNYRHVKTMPVGDAVLDTFMVNLTLIVENI 593
Qy 655 DFKVIELYTRDEKRLSNVDFIETMFREYNYAQRVSGRLKDLDDLSTNRNQFVDAFGSLM 714
Db 584 DFOVIELYSREBKRMSTAFDIETMFREYNYTQRTVGLRRDLTDLATNRNQFVDAFGSLM 643
Qy 715 DDLGAGVQTVVNAVSGVATLFSSIVTGFINFKNPGGMLMIIVJGVLFPAIYFLTKTK 774
Db 644 DDLGVVGVKTVLNAVSGVATLFSSIVGFINFKNPGGMLLFLGLIAAVVITVILLNRKAK 703
Qy 775 IYETAPIKMIYPIIDKLKEREKSEIAPISEBELRIVLAMHIHQONSHMETK----- 827
Db 704 RFAQNPVQMI-PDIKTITTSQREELQVDPISKHELDRLIMLAMHDYHASKQPEKQDEEQS 762
Qy 828 TRKDPKDSILTRAQNMLRKESGYSLNKAESVE 860
Db 763 TTSQPAD--LNKAKNVLRRAAGYKPLKRTDSFE 793

RESULT 3
US-10-055-364-45
; Sequence 45, Application US/10055364
; Publication No. US20020155433A1
; GENERAL INFORMATION:
; APPLICANT: Patience, Clive
```

```
/ TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use
/ FILE REFERENCE: 61750-379
/ CURRENT APPLICATION NUMBER: US/10/055,364
/ CURRENT FILING DATE: 2002-01-23
/ PRIOR APPLICATION NUMBER: US/09/612,204
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: US/60/142,736
/ PRIOR FILING DATE: 1999-07-08
/ PRIOR APPLICATION NUMBER: US/60/168,532
/ PRIOR FILING DATE: 1999-12-02
/ NUMBER OF SEQ ID NOS: 55
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 45
/ LENGTH: 830
/ TYPE: PRF
/ ORGANISM: Alcelaphine herpesvirus
US-10-055-364-45
```

```
Query Match      53.6%; Score 2438; DB 14; Length 830;
Best Local Similarity 57.4%; Pred. No. 1e-199;
Matches 459; Conservative 129; Mismatches 166; Indels 46; Gaps 7;

Qy      64 NENRDESEQN---KNYSGSPFPYRVCSASGVDFRQTDHVCPCDASDMVHSEGIILLI 120
Db      42 SDNIIRQNNNTAKGIHSDPSAFPFRVCSASNIGDIFRQTSHPSCNPKDKEHNEGILLI 101

Qy      121 YKONIIPFMRVRKRVKVVTTSTVYNGIYSDSITNOHTFYKSIPEWETEKMDTIYQCFNS 180
Db      102 FKENIVPVYKRVKRVKVVTTSTVYNGIYADAVTNOHVFSKVPYIETRRMDTIYQCFNS 161

Qy      181 LRLNTGGNLLTVDRDDINMTVFLQVPDGVTPDKRYGSOPELYLEPGWFGSVRRRTTV 240
Db      162 LDVTVGGNLLVYTDNGSNMTVDLPQDGLNSVRYHSQPEIHAEPGLGGLYRRRTTV 221

Qy      241 NCELMDMFARSNPPDFFWTATGDTVEMSPFW-----SGDDHKNKHEKPFVFSV 291
Db      222 NCEVTEDARAVPPRYEITNIGDTIEMSPWKAWNETEFSGEPRD-----LTV 272

Qy      292 INNYKVVDYQNRGTVPGLKTRIFLDREEVYTLSEKHLKNMSYCPILTAKAFVNGIQTEHS 351
Db      273 AKDYRVVDYKFRGTQPGHTRIFVDKEEYTLSSAQOFRNISYCRWAHWKSDNAIKTBHG 332

Qy      352 GSYHFVANDITASFTTSKEDKEFNNTYHCLNEEIKAEIEKKYAKVNSTHYSKYGDLKYFK 411
Db      333 KSLHFVANDITASFTYPTNTQREVLGKHVCLNNTIESELKSLAKVNDTHSPNGTAQYVL 392

Qy      412 TDGGLYLVWQPLIQNRLLDAKNKLN-----NETYSRRSRQAESTTDPMMEMTONG 462
Db      393 TNGGLLVWQPLVQOQKLLDAKGLLDVAKKQQTNTTTTTRSRQRORSVS-----SG 444

Qy      463 AGGEYSSENSITVAQVQYADNLRIRINNILEDLSKAWCRQRAALVMNELSKINPTSV 522
Db      445 IDDVYTAESTILLTQIQFAYDTLRAQINNVLIELSRACRQRAALVMNELSKINPTSV 504

Qy      523 MSMIYNRPVSAKRIGDVISVSNCCIYVDQTSVSLHKSRLLSA-SDEKCFSRPPVTFKPMN 581
Db      505 MSSYGRPVSAKRIGDVISVSHCVVDQDSVSLHSMRVPGRDKTHECYSRPPVTFKFIN 564

Qy      582 DSTIYKQGLGVNNEILLTYYLETQENTEYFQAKTOMYIKYVNHUKTVPLSSITLTD 641
Db      565 DSHLYKQGLGVNNEILLTATTAVEICHEENTHYFQGGNNMYFYKNYRHKVTPVGDVATLD 624

Qy      642 TPIALNFTLLENVDFKVELYTRDKSLNSVPDIETMERENYIAQRYSGRKLKOLLDLST 701
Db      625 TFWUNLTLVENIDFQVIELYSREBKRSTAFDIETMERENYIYQRYVTGLRDLTDLAT 684

Qy      702 NRNQVDFAGSLMDDLGLVAGVGTNVAVSGVATLFSIVTGFNFTKPNFPGGLMILVIG 761
Db      685 NRNQVDFAGSLMDDLGLVGVGTVLNAVSVATLFSIVSGINFTKPNFPGGLMFLGLIAA 744

Qy      762 VLFALYFTTKTKIVETAPIKMIYPEIDKLKEREKSEIAPISEBEELERIVLAMHIHOON 821
Db      745 VVITVILLNRKAKRAQNPVQMIVPDIKTITSQREELQVDPISKHELDRIIMLAHMDYHAS 804
```

```
Qy      822 SHMETK-----TRKDPKD 834
Db      805 KPESKQDEEQSGTTSQPAD 824
```

RESULT 4

```
US-10-055-364-43
/ Sequence 43, Application US/10055364
/ Publication No. US2002015543A1
/ GENERAL INFORMATION:
/ APPLICANT: Patience, Clive
/ TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use
/ FILE REFERENCE: 61750-379
/ CURRENT APPLICATION NUMBER: US/10/055,364
/ CURRENT FILING DATE: 2002-01-23
/ PRIOR APPLICATION NUMBER: US/09/612,204
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: US/60/142,736
/ PRIOR FILING DATE: 1999-07-08
/ PRIOR APPLICATION NUMBER: US/60/168,532
/ PRIOR FILING DATE: 1999-12-02
/ NUMBER OF SEQ ID NOS: 55
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 43
/ LENGTH: 824
/ TYPE: PRF
/ ORGANISM: Eguine herpesvirus 2
US-10-055-364-43
```

```
Query Match      42.7%; Score 1940; DB 14; Length 824;
Best Local Similarity 47.0%; Pred. No. 4.9e-157;
Matches 395; Conservative 141; Mismatches 241; Indels 64; Gaps 16;
```

```
Qy      11 VLALWLYQVALYSLSTAETGVTSPNPTATWSTESPTUHYGTHDSHGREGNNENRDS 70
Db      10 VLCLWCV--AALLCQGVAEVVA-----ETTP-----FATH-----RPEVVAE 46

Qy      71 EQNKNIYGSSTFPYRVCSASGV-GDYFRQTDHVCPCDASDMVHSEGIILLIYKQNIIPFM 129
Db      47 ENPANPF-----LPRVCGASPTGGEIIFRPLEBSCENTEDKHIEGIALIYKNIIPVY 101

Qy      130 FRVKYRKVVTTSTVYNGIYSDSITNOHTFYKSIPEWETEKMDTIYOCFNSRLNTGNNL 189
Db      102 FNVKRYKIMTSITYKGSWSDAITNOHTRSYAVLYEVQMDHYQCFSAVQVNEGHHV 161

Qy      190 LTVYDRDDINMTVFLQVPDGVTPDKRYGSOPELYLEPGWFGSVRRRTTVNCELMDMFA 249
Db      162 NTYYDRDGMNETAFKPKADGLTSSITRYQSOPEVYATPRNLLWSYTRTTTVCNCEVTEMSA 221

Qy      250 RSNPPDPFVATGDTVEMSPFWSGEDDHENKHEKPFVSVINNYKVVDYQNRGTVPGL 309
Db      222 RSMKPFPEFFVTSGVDTIEMSPFLKENGTEPEKILKRPHSIQLLKNYAVTKY---GVGLG 277

Qy      310 K---TFILDRBEYTLSSWKHLKNMSYCPILTAKAFYNGIQTEHSGSYHFPVANDITAS 365
Db      278 QADNATFFPAIFDYSLSWKATTESSYCDLLWKGSNAIQTHNSLSLFIANDITAS 337

Qy      366 TTSKEDKEFNNTYHCLNEEIKAEIEKKYAKVNSTHYSKYGDLKYFKTDGGLYLVWQPLIQ 425
Db      338 STPLEEENFNETFKCIWNNTQBEIQKKLKEVEKTRPNGTAKVYKTTGNLYLVWQPLIQ 397

Qy      426 NRLLDAKNKLNN-----ETYSRRSRQAESTTDPMMEMTONGAGGEYSS-ENS 472
Db      398 IDLLDTHAKLYNLNTNATASPTSTPTTSPRRRRDTS-----VSGGNGNNGDNSTKEBS 450

Qy      473 ITVAQVQYADNLRIRINNILEDLSKAWCRQRAALVMNELSKINPTSVMSTYNNRPVS 532
Db      451 VASQVQFAYDNLRLKSRINRVLGSLRAWCRQRAALVMNELSKINPTSVMSTYNNRPVS 510

Qy      533 AKRIGDVISVSNCCIYVDQTSVSLHKSRLLSASDEKCFSRPPVTFKPMNDSTIYKQGLV 592
Db      511 AKLIGDVVSDCISVDQKSVFVHKNMK-VPGKEDLCYTRPVVGVFKFINGSELPAGQLGP 569
```

Qy	593	NNEILLTYYLETCQENTYYFQAKTDMYIYKVEYHUKTVPLSSITTLDTOTFIALNPTLLE	652
Db	570	RNEIVLSTSQVEVCQHSCEHYFQAGNQMYKYKDYVVSTLNLTDITPLHTMITLNLSLVE	629
Qy	653	NVDFKVIELYTRDEXKRLSNVFDIETMPREYNYYAQRVSGLRKOLL-D-LSTNRNQFVDAFG	711
Db	630	NIDFKVIELYSKTEKRLSNVFDIETMPREYNYYTQNLNGLRKOLLDSIDHGRDSFIQTGLG	689
Qy	712	SLMDDLGAQGTVVNAVSGVATLFSISVITGFINFKNPFQGMILMIIVI-GVLFAIYFLT	770
Db	690	DIMQDLGTGKGVVNAVSGVSLFGSIVSGVISPFKNPFGGMILLIILIIAGVVVVVYLFMT	749
Qy	771	KXTKIYETAPIKWIPEIDKLKEREGKSEIAPISEEELEIRIVLAMHTHOQNSHMETKTRK	830
Db	750	RRSIY-SAPIRMLIPGVERRAAQEPGAH---PVSEDIQRNMLMGMHQFQORQRAEEEAR	805
Qy	831	D 831	
Db	806	E 806	
RESULT 5			
US-10-055-364-37			
; Sequence 37, Application US/10055364			
; Publication No. US20020155433A1			
; GENERAL INFORMATION:			
; APPLICANT: Patience, Clive			
; TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use			
; FILE REFERENCE: 61750-379			
; CURRENT APPLICATION NUMBER: US/10/055,364			
; CURRENT FILING DATE: 2002-01-23			
; PRIOR APPLICATION NUMBER: US/09/612,204			
; PRIOR FILING DATE: 2000-07-07			
; PRIOR APPLICATION NUMBER: US/60/142,736			
; PRIOR FILING DATE: 1999-07-08			
; PRIOR APPLICATION NUMBER: US/60/168,532			
; PRIOR FILING DATE: 1999-12-02			
; NUMBER OF SEQ ID NOS: 55			
; SOFTWARE: PatentIn version 3.0			
; SEQ ID NO 37			
; LENGTH: 823			
; TYPE: PRT			
; ORGANISM: Human herpesvirus 8			
US-10-055-364-37			

Query Match	42.2%	Score 1918;	DB 14;	Length 823;
Best Local Similarity	48.5%;	Pred. No. 3.7e-155;		
Matches	378;	Conservative 134;	Mismatches 233;	Indels 34; Gaps 10;
QY	83	PPYRVCASGVGVFRQTDHVHCPDASDMVHSGEILLIYKNIIPFMFRVKRYKVVTTS	142	
DB	63	YQFRVCASGATGELFRFNLEQTCDDTKDKYHQEGILLVYKKNIIPVHFVKRRYRKIATSV	122	
QY	143	TVYNGIYSDITNOHTYKYSIEPWETEKMDTIYQCFNSLRINTCGNLLTVYDRDDIINMTV	202	
DB	123	TVYRGLTSAITNKYELPRPVLYEISHMGSTYQCFSSMKVNVGVNVENTFTDRDDVNTTV	182	
QY	203	FLOPVDGVTDPVKRYGSOPELYLEPGHFWGYSRRRTTVNCELMDMFARSNPPDFVFYAT	262	
DB	183	FLQPEVGLTDNIQRYFSQPIYAEPPGFPGIYRVRTTVNCEIYVDMIARSAEPYNYFTSL	242	
QY	263	GDTVEMSPFFWSGEDDHENKMKHPWF-VSVINNYKVVDYQNRGTVPGLKTRIFLDRREYT	321	
DB	243	GDTVEVSPFCVNESSCSTTTPSNKGLSVQVVLNHTVVTVSDRGTSPTPQNRIFVETGAYT	302	
QY	322	LSWEKHLKNMYSCLPTLTKWAFYNGIQTEHSGSYHFVANDITASPTTSKEDMKENPTYHC	381	
DB	303	LSWASESKTTAVCPALAWKTFPRSIQTHEDSFHFVANEIATPTAPLTPVANFTDYSC	362	
QY	382	LNBEIKAEIKKYAKVNSTHSKYGDLYFKTDDGGLYLVWQPLIQNRLLDAAKNLNNET--	439	
DB	363	LTSDIMTLNASKAKLSTAVPNCVTQYFHTTGLSLYLVWQPMASAINITHAGDSGNPTSS	422	

```

Qy 440 -----YRRRRRQAESTTDDPMWMTGNGAGGEYSSENSITVAQVQVAYVNLRR 488
Db 423 PPSASPMWTTSSASRRKRRSASTA-----AAGGGGSTDN-LSYTLQQLFAYDKLRDG 471
Qy 489 INNITLEDLKWACREOHRRAALVNNELSKINPTSVMSTMYNRPVSAKRIGDVISVNCIVV 548
Db 472 INQVLEELSRACREOVRDNLMMWYELSKINPTSVMTAIYRPPVSAKFVGDAISVTECINV 531
Qy 549 DOTSVLHKSRLLSASDEKCFSPVPVTFKPMNDSTYKKGQGVNNEILLTITVYLETQOE 608
Db 532 DQSVNTHKSLR--TNSKDVCIARPLVTFKFLNSNLTFTGQLGARNEILLTNQVETCKD 589
Qy 609 NTEYFQAKDMYIKNYEHLKTVPLSLSTITLDTFIALNFTLLENVDFKVIELYTRDEKR 668
Db 590 TCEHYFITRNETLVKYDVAYLRTINTTIDISTLNTFIALNLFSIQNIDPKAIELYSSAEKR 649
Qy 669 L-SNVPDIETMFRYNNYAAOVSGLRKDLLD-LSTNRNQFVDATGSLMDDLGAVGQTVVN 726
Db 650 LASSVFDLETMFRYNNYTHRLAGUREDLONTIDMNERFVRDLSEIVADLGGIGKTVVN 709
Qy 727 AVSGVATLFSISVTFGNFTKNPFGGMMLIIVGLVFAIYFLTKTKKIYETAPIKMIYP 786
Db 710 VASSVVTLCGLSVTFGNFTKHPGLGGMLIIVIAIILIEPMLSRRTNTTIAQAPVKMIYP 769
Qy 787 EIDKLKREGKSETAPISEEBELERIVLAHMHQONSHMETKTRDKPKDSILTRAQNMLR 845
Db 770 DVDR--RAPPSGAP-TREEIKNILLQMHQLOQERQKADDLKK-S-TPSVFQRTANGLR 823

RESULT 6
US-10-055-364-42
; Sequence 42, Application US/10055364
; Publication No. US2002015543A1
; GENERAL INFORMATION:
; APPLICANT: Patience, Clive
; TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use
; FILE INVENTION: 61750-379
; CURRENT APPLICATION NUMBER: US/10/055,364
; CURRENT FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: US/09/612,204
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US/60/142,736
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: US/60/168,532
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 42
; LENGTH: 792
; TYPE: PRT
; ORGANISM: Herpesvirus saimiri
US-10-055-364-42

```

	Query Match	42.1%	Score 1916;	DB 14;	Length 792;
	Best Local Similarity	46.9%	Pred. No. 5.2e-155;		
	Matches 366;	Conservative 155;	Mismatches 228;	Indels 32;	Gaps 9;
Qy	71	EQKNIVGSPSTPPYRCSAGSGDVFRFOTDHVCPDASDMVHSEGILLIVYKONIIIPMF	130		
Db	28	EKNQTAIQEYKYRVCASATGGELFRFDLDRCTSTEDKVHKEGILLIVYKNIVYIP	87		
Qy	131	RVRYRKVVTSTVYNGIYSD--SITNQHTFYKSIIEPWETEKMDTIVQCFNSLRNTGQN	189		
Db	88	KVRRYKITTSVRFNGTWRGVAITKNWLSRAVPKYEIDIMDKTYQCHNCMQIEVNGM	147		
Qy	189	LLTVYDRDDNNMTVFLQDPGVTPDKVRGSGQPELYLPCGFWGSRVRRRTTVNCELMDMF	248		
Db	148	LNSYIEDRGNNKVVDLRFVLDGLTGATRYISQPKVPADPGLMGVTRTRTVNCEIIVDMF	207		
Qy	249	ARSNPPDFEVTATGDTVENSFPWMSGEDDHENKKHKPWFVSVINNYKVVDYQNRGTVP	308		
Db	208	ARSADPTTFVTALGDTVEVSPFCDDVNSCPNATDVL--VOIDLNTHTVVDYGNRATSOQ	265		


```
QY 309 GKTRIFLDREBYTLSEKHLKNNMSCPLTLWKAFYNGIQTEHSGSYHFVANDITASFTTS 368
Db 266 HKKRIFAHTLDYSVSWEAVNKSASVCSVMFWKSFQRAIQTEHDLIYHFIANEITAGFSTV 325
QY 369 KEDMKEFNTTYCLNEEIKAEIEKKYAKVNSTHSGYKGLKPKTKDGLYLVNQPLIQNRL 428
Db 326 KEPLANFTSDYNCLMTHINTTLEDKIARVNNHTPNGTAEYQTEGGMLVWQPLIAIEL 385
QY 429 LQAKNKLNE-----TYSRERSRQAESTTTPMMEMTNGAGGEYSSENSITVAQVOYA 481
Db 386 EAMLEAATSPVTPSAPTSRSSKRAIRSIRDV-----SAG-----SENNVFLSQIOYA 434
QY 482 YDNLRIRINNILEDLSKAWCRQHRRAALVWNLSEKINPTSMVSMIYNRPVSAKRICDVIS 541
Db 435 YDKLRQSINNVLLEALITWCREQVRQTMVWEIAKINPTSVMTAIYKRPVSRKALGDVIS 494
QY 542 VENCIVVDOTSVLHSLKLLLSASDEKCFSPRPVTFKFWNDSTIYKQGLGVNNEILLTTT 601
Db 495 VTECINVDOSSVSIHKSLLK--TENNDICYSRPPVTFKFNVSQLFKQGLGARNELLSES 552
QY 602 YLETCOENTYVFOAKTDMYIYKNYEHKLTVPSSITLTDTFIALNFTLLENVDFKVEL 661
Db 553 LVENCHONAETFTAKNETHFNKNVHVETLFPVNNISLTDTFLALNFTFIENIDFRAVEL 612
QY 662 YTRDEKRLSNVFDIETMFREYNYAQRVSGRLKLLDLS--TNRNQFVDAFGSLMDDLGA 719
Db 613 YSSGERKLANVFDLEMTFREYNYAQSISGLRKD--FDSQRNRRDRIIQDFSEIILADLGS 671
QY 720 VQGTVVNAVSGVATLFFSSIVTGFINFKNPFQGLMLMIIVIGVLFALVPLTKTKIYETA 779
Db 672 IGVIVNVASGAFSLFGGIVTGILNFKNPLGCMFTLLIGAVIILVILLVRRTNNMSQA 731
QY 780 PLKMIYPEIDKLKEREGSEIAPISEEELERIVLAWHHHQONSHMETKTRKDPKDSILTR 839
Db 732 PIRMIYDVEK-----SKSTVTPMBEPETIKQILGHHNNQOEAYKKKEQRAARPSIFRQ 786
QY 840 A 840
Db 787 A 787
```

RESULT 7

```
US-10-055-364-38
; Sequence 38, Application US/10055364
; Publication No. US20020155433A1
```

GENERAL INFORMATION:

```
; APPLICANT: Patience, Clive
; TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use
```

```
; FILE REFERENCE: 61750-379
```

```
; CURRENT APPLICATION NUMBER: US/10/055,364
```

```
; CURRENT FILING DATE: 2002-01-23
```

```
; PRIOR APPLICATION NUMBER: US/09/612,204
```

```
; PRIOR FILING DATE: 2000-07-07
```

```
; PRIOR APPLICATION NUMBER: US/60/142,736
```

```
; PRIOR FILING DATE: 1999-07-08
```

```
; PRIOR APPLICATION NUMBER: US/60/168,532
```

```
; PRIOR FILING DATE: 1999-12-02
```

```
; NUMBER OF SEQ ID NOS: 55
```

```
; SOFTWARE: Patent in version 3.0
```

```
; SEQ ID NO 38
```

```
; LENGTH: 808
```

```
; TYPE: PRT
```

```
; ORGANISM: Rhesus monkey rhadinovirus
```

```
US-10-055-364-38
```

Query Match

```
Best Local Similarity 41.8%; Score 1902; DB 14; Length 808;
```

```
Matches 384; Conservative 142; Mismatches 240; Indels 102; Gaps 14;
```

```
QY 11 VIALMWLYXVALYSIAETGVTSPNTATWSTESPLTGHYTHDSHGGRGNENRDSR 70
```

```
Db 10 LLRAWV--IIAIGTAVGE-NVTTPKGATTAKTP----- 42
```

```
QY 71 EONKNIYGPST-----FPIRVCSASGVGVFRFQTDHVCPCDASDMVHSEGIL 118
Db 43 -----GPSTPTPPENPPAEAFKFRVCSASATGELFRNLEKTCPTGTEKTHQEGIL 94
QY 119 LIYKONIIPIPMFRVRYKRVKVVTTSTVYNGIYSDSITNQHTFYKSIEMETSEKMDTIYQCF 178
Db 95 MVFKKNIYVPHIFKVRRYRKVATSIVTAVGTWGTETATVTKQEVIRVPVQVEIHNMDTTYQCF 154
QY 179 NSLRLNTGGNLLTYVDRDDINMTVFLQVGDVTPDVKRYGSOPELYLEPGHFWGYSYRRT 238
Db 155 SSMRVNNGVINYNTYDRDFTNQTVLQPVVEGLTNIQIRYFSQPLVLYTTPGMPFGIYRVT 214
QY 239 TVNCELMDMFARSNPPDFDFFVATGDTVEMSPFWSGE-----DDHENKMKHEKFWFSVI 292
Db 215 TVNCEIVDMIARSAPYSYFVTALGDIVEVSPFCHNDSTCSVAEKTENGLCAR-----VL 269
QY 293 NNYKVVDYQNRGTVPVLGKTRIFLDRBEYTLSEKHLKNNMSCYCLTLWKAFYNGIQTEHSG 352
Db 270 TNYTMVDPATR--APTETRVFADSGEYTVSWKAEDPKSAVCALTLWKTFPRAIQTTHA 327
QY 353 SYHFVANDITASFTTSKEDMKFNTTYHCLNEEIKAEIEKKYAKVNSTHSGYKGLKPKTK 412
Db 328 SYHFVANDVATFTSPUSEVANFTGTYSCLDEVIQKTLNDTIKKLSOETHVTNGSAQYTKT 387
QY 413 DGGLYLVWQPLIQNRLLDKAKNNNETYS-----RRSRQAESTTTPMMEMTGN 461
Db 388 EGGFLWQPLTLPLSLVDENGLNGTTPAPATTSTVSRVRSVNTNEQATD----- 439
QY 462 GAGGEYSSENSITVAQVQYAYDNIRIRINNILEDLSKAWCREQHRRAALVWNLSEKINPTS 521
Db 440 -----NLAAPQLQFAYDKLRASINKVLEELSRACREQVRDTYMMYELSKINPTS 489
QY 522 VMSMIYNRPVSADIGDVISVNCIVVDOTSVLHSLKLLLSASDEKCFSPRPVTFKPMN 581
Db 490 VMTAIYGRPSVSAKFGVDAISVTDCVAVDQASVSIHKSLLR--TSTPGMCYSRPPVTFRFLN 547
QY 582 DSTIYKQGLGVNNEILLTYYTLETCOENTYVFOAKTDMYIYKNYEHKLTVPSSITLTD 641
Db 548 STTLFKQGLGPRNEIILTDNQVEACKETCEHYFTASNVTYVYKDYVFKKINTSEISLGL 607
QY 642 TFIALNFTLLENVDPKVIELYTRDEKRLS--NVFDIETMFREYNYAQRVSGRLKDL---L 697
Db 608 TFIALNLSFIENIDFRVIELYSRAEKKLSGVFDIETMFREYNYVQRLAGLREDLDNTI 667
QY 698 DLSTNRNQFVDAFGSLMDDLCAVGQTVVNAVSGVATLFFSSIVTGFINFKNPFQGLMLMI 757
Db 668 DL--NRDLARDLSEIVADLGDVGRTVNVVASSVITLFGSIVSGFINFIKSPFGGLMIL 725
QY 758 VVIGVLFALVPLTKTKIYETAPIKMIYPEIDKLKEREGSEIAPISEEELERIVLAWHI 817
Db 726 VIVAVVLIVFALNRRTNAIAQAPIRMIMYDPIDKMQPSGGK-----VDOEQIKNILAGMHQ 780
QY 818 HQONSHMETKTRKDPKDSILTRAQNMRL 845
Db 781 LQEEERRRLDEQORSAPSLFRASDGLK 808
```

RESULT 8

```
US-10-055-364-41
```

```
; Sequence 41, Application US/10055364
```

```
; Publication No. US20020155433A1
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: Patience, Clive
```

```
; TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use
```

```
; FILE REFERENCE: 61750-379
```

```
; CURRENT APPLICATION NUMBER: US/10/055,364
```

```
; CURRENT FILING DATE: 2002-01-23
```

```
; PRIOR APPLICATION NUMBER: US/09/612,204
```

```
; PRIOR FILING DATE: 2000-07-07
```

```
; PRIOR APPLICATION NUMBER: US/60/142,736
```

```
; PRIOR FILING DATE: 1999-07-08
```

```
; PRIOR APPLICATION NUMBER: US/60/168,532
```

```
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 41
; LENGTH: 791
; TYPE: PRT
; ORGANISM: Ateline herpesvirus 3
US-10-055-364-41

Query Match      41.8%; Score 1898.5; DB 14; Length 791;
Best Local Similarity 46.6%; Pred. No. 1.6e-153;
Matches 367; Conservative 144; Mismatches 235; Indels 41; Gaps 10;

Qy 69 SEQKNYIGSPSPFYRVCASGVDVRFOTDHCPCDASDMVHSEGLILYKQNIIPF 128
Db 26 SDENKTPAIEKEFKYRVCASASTGELFRNLDRACPSTEDKVRHREGILLVYKKNIVPH 85

Qy 129 MFRYKRVKVVTTSTVNGIYSD--SITNQHFTFKYSIEPMEWKMDTIYQCFNSRLNTG 186
Db 86 IFKVRVYKKIATSVRIENGWSREGVAITNKWELGRAVPKYEINLMDKNYQCHNCMQIEVN 145

Qy 187 GNLLTYVDRDDINMTVFLQPDVGVTPDVKRYGSOPELYLEPGWFGSYRRRTTVCNCELM 246
Db 146 GLLENSYCDRDGNKNTVDLKPVDGUTGAI TRYVSQPKIFADAGMLWGTYKTRTTVNCNCE 205

Qy 247 MFARSNPPDFVTATGDTVEMSPFWSGEDDHENKHEKMPFVGVIN----NYKVVDYQ 301
Db 206 MFARSADPYTFVVALGDTVEVSPFCDANSCPNAS-----DVLSSQVDFNHTVVVDYG 258

Qy 302 NRGTVPLGKTRIFLDREYTLWSKHLKNMSYCPLLTWKAFYNGIQTEHSGSYHFVANDI 361
Db 259 NRATSOQHGKRIFAHTLDYSVWEAINKTTSCVSMVFWKGFQRAIQTEHDSYHFIANEI 318

Qy 362 TASTTTSKEDMKENTYHCHNEBIKAEIEKKYAKVNSTHSKYGDLKYFTDGLGLYLWQ 421
Db 319 TAGESTSKETLASFSSEYSCILMSDINSTLDKIGRVNNTVHPNGTAQYFKTEGGMILWQ 378

Qy 422 PLIQNRLLDANKLNN-----EYSSRRRQAEISTTDPMMEMTGNCGAGGEYSSENSITV 475
Db 379 PLTAIELEAMEIATTVSPPLSTAHLTSTRGKRKDV-----SAG-----SENSVLL 427

Qy 476 AQOYAYDNLRIINNILEDLSKAWCREQRAALVNNELSKINPTSVMSMIYRNPVSAGR 535
Db 428 AQIQYAYDKLRQSNVLEELAITWCREQVQRTMIWYIEAKINPTSVMTAIYKPVSAKA 487

Qy 536 IGDVISVNCIVDQTSVLSHKSURLLSASDEKCFSPPTVTFKPMNDSTIYKQGLGVNNE 595
Db 488 LGDVISVTECINVDQTSVSIHKSILK--TTNNDVCYSRPPVTFKPVNSSQLFKGQGLGARNE 545

Qy 596 ILLFTTYLETQCOENTYFYFOAKTDMYLYKNVEHLKTVPLSSITTLDTFIALNFTLENVD 655
Db 546 ILLSESIVNCHQNAEHFFAKNETHYFNKNVLAHVELTPLNTISITLDTFLALNFTLEND 605

Qy 656 FKVIELYTRDEKRLSNVFDIETMFREYNYYAQRSGLRKDLLDS--TNRNQFVDFAGSL 713
Db 606 FKAVELYSSEGERKLANVFDLETMFREYNYYAQRSISGLRKD--FDNSQRNRRDRIIQDFSEI 664

Qy 714 MDDLCAVGQTVNAVSGVATLFPSSIVTGFINFKNPGGMLMIIVIGVLFAIYFLTKYT 773
Db 665 LADIGSGIKGVINVIASSAFSLFGGIVTGILNFINKNPLGGMILTFLLVGAIIILVILVRRT 724

Qy 774 KIVETAPIKMIYPIIDKLKEREGKSEIAPISSEBELRIVLAMHIHOONSHMETKTRDKP 833
Db 725 NNNQSAQIRMIYDPDIEK-----SRSSVTPTPEPEVIKQILLOMNMNQOEEYKKEEHKASQ 779

Qy 834 DSILTRA 840
Db 780 PSFLKRA 786

RESULT 9
US-10-055-364-44
; Sequence 44, Application US/10055364
```

```
; Publication No. US20020155433A1
; GENERAL INFORMATION:
; APPLICANT: Patience, Clive
; TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use
; FILE REFERENCE: 61750-379
; CURRENT APPLICATION NUMBER: US/10/055,364
; CURRENT FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: US/09/612,204
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US/60/142,736
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: US/60/168,532
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 44
; LENGTH: 818
; TYPE: PRT
; ORGANISM: Equine herpesvirus 5
US-10-055-364-44

Query Match      41.2%; Score 1871.5; DB 14; Length 818;
Best Local Similarity 46.2%; Pred. No. 3.6e-151;
Matches 368; Conservative 149; Mismatches 223; Indels 57; Gaps 13;

Qy 85 YRYCASGVDVRFOTDHCPCDASDMVHSEGLILYKQNIIPFMRVRYKRVKVVTTSTV 144
Db 56 FRVCGASPGTEIFRFPLEENCPNTEDEKHEVGEILLIYKTNIVPIYFNVRKYRKLVTSTI 115

Qy 145 YNGIYSDSITNQHFTFKYSIEPMEWKMDTIYQCFNSRLNTGGLLLTYVDRDDINMTVFL 204
Db 116 YKWSQDAITNQYTSSFAMPLMEARLVDYNYECYNGIQTEHGLTYVDRDGYNESVRL 175

Qy 205 QPVDGVTPDVKRYGSOPELYLEPGWFGSYRRRTTVCNCELMFARSNPPDFVTATGD 264
Db 176 VPADGLTTSIRRYVHSQPELYVTPRNLLMSYTRTTVNCVCEVIDMTARSHKPPFEYFVTASGD 235

Qy 265 TVEMSPFWSGEDDHENKHEKMPFVGVINNYKVVDYQNRGCTVPLGK----TRIFLOREEY 320
Db 236 SIETSPFYT-----NASRRVP--VQVLYNSVTDY-----GVGLSGSENVTRFATLNDP 283

Qy 321 TLSWEKHLKNMSYCPLLTWKAFYNGIQTEHSGSYHFVANDITASFTTS--KEDMKENTTY 379
Db 284 SISKKAATENSSYCPVLVWKGPSPAIQTKHESYHFIADAVTASFTTPTLTDTSYENTTY 343

Qy 380 HCLNEBIKAEIEKKYAKVNSTHSKYGDLKYFTDGLGLYLWQPLIQNRLLDANKLNN-- 437
Db 344 QCAWQDIEGEIQKRFDPVSKTHARNGSVQIYKTSGNLYVWQPLVQLDLLAAHAKTINST 403

Qy 438 -----EYSSRRRQAEISTTDPMMEMTGNCGAGGEYSSENSITVAOQYAYD 483
Db 404 DNSTSTTAPNTTTSRRKRDRDTGNTA-----TNSSNNSSMBEELATSSQVQFAYD 457

Qy 484 NLRIRINNILEDLSKAWCREQRAALVNNELSKINPTSVMSMIYRNPVSAGRIGDVISVS 543
Db 458 QLRKSNRVLEQLSRVWCQYRASLMWYELSKINPTSVMSAIYGRPVSAKLGVDDVQVIS 517

Qy 544 NCIVDQTSVLSHKSURLLSASDEKCFSPPTVTFKPMNDSTIYKQGLGVNNEILLTTTYL 603
Db 518 DCITVDQESVVFVHRNLR--VPGSKDLCTYTRPVGVGKFIENGSELPVQGLGARNEILLSTNLV 576

Qy 604 ETCOENTYFYFOAKTDMYLYKNVEHLKTVPLSSITTLDTFIALNFTLENVDKVLLEYLT 663
Db 577 EVCQHSCEHYFOGNNHIYKKNVEYVSTWNLTVDVPTLHTMTITLNLSENVDFQVQLQYS 636

Qy 664 RDEKRLSNVFDIETMFREYNYYAQRSGLRKDLLD--LSTNRNQFVDAFGLSMDLDLGAVGQ 722
Db 637 QREKRLSNVFDIETMFREYNYYTQNLKGLRKDLLDSDSIHGRDSFIQFLGDLVQDLVPVGD 696

Qy 723 TVNAVSGVATLFPSSIVTGFINFKNPGGMLMIIVIGVLFAIY--FLTKTKKIYETAPI 781
Db 697 VIVNVASGVFLSGSIVSGVISFLKNPLGAILTIALIVGGIIVLYLFIITRSRTVYQ-API 755
```

```
QY 782 KMIYPIDKLKEREGKSEIAPISEELERIVLAMHIHQNSHMETKTRKDPKDSILTRAQ 841
Db 756 RMLYPEVD---RAPQNVQPIPEDQVRGILLAMHQFQQQQQQQQQQQ-----E 801

QY 842 NMLRKESGYSLNKAES 858
Db 802 EHTORRSIFDTIRESTS 818

RESULT 10
US-10-055-364-46
; Sequence 46, Application US/10055364
; Publication No. US20020155433A1
; GENERAL INFORMATION:
; APPLICANT: Patience, Clive
; TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use
; FILE REFERENCE: 61750-379
; CURRENT APPLICATION NUMBER: US/10/055,364
; CURRENT FILING DATE: 2002-01-23.
; PRIOR APPLICATION NUMBER: US/09/612,204
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US/60/142,736
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: US/60/168,532
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 46
; LENGTH: 829
; TYPE: PRT
; ORGANISM: Epstein-Barr virus
US-10-055-364-46

Query Match 40.6%; Score 1844.5; DB 14; Length 829;
Best Local Similarity 46.1%; Pred. No. 7.5e-149;
Matches 367; Conservative 148; Mismatches 240; Indels 41; Gaps 12;

QY 81 STFPYRVCSAGVGDFRQTDHVCDD-ASDMVHSEGLLIYKQNIIPFMFRVKRYKVV 139
Db 44 TSFPFRVCSLGHDLFRSSDIQCPSTRENHTREGLLMVFKDNIIPYFVKRSTYKIV 103

QY 140 TTSTVYNGIYSINQHTFYKSIPEWETEKMDTIYQCFNSRLNTGNNLLTYVDRDIN 199
Db 104 TNLINYNGIYADSVNRHEEKESVSDYETDQDITTYQCVNAVMTKDGLTRVYVDRDGN 163

QY 200 MVELQPDVGVDPDKRYGSOPELYLEPCGFWGVSRRRTTNCCLMDMFARSNPFPDFV 259
Db 164 ITVNLKPTGGLANGVRRYASQTELYDAPGGLIWTYTRTTVNCCLITDMMAKSNPFPDFV 223

QY 260 TATGDTVEMSPWSEDDHENKMKHPWFSVINNYKVVDYQNRGTVPGLKTRIFLDREE 319
Db 224 TTTGQTVEMSPYDCK--NKETFEHRAFSFHYRTNYKIVDYNRGTNFGERRAFLDKGT 281

QY 320 YTLSEKHLKN--MSVCPLTLKAFYNGIQTEHSGSYHFVANDITASFTTSKEDMKFNTT 378
Db 282 YTLNWK--LENRTAYCPLQHMGTQSTIATETGKSIHFVTDGTSFVNTTVTGIELPDA 339

QY 379 YHCLNEEIKAEIEKKYAKVNSTHSHKYG--LKYFKTDGGLYLVQPLIQNRLLDKXN--KL 435
Db 340 FKCIIEQVKNTHKEYEAQVDYRTKGQEAITYFITSGGLLAWLPLTPRSLATVKNLTEL 399

QY 436 NNETYSS-----RRSRQAEISTTDPMMWMTGNGAGEVSENSI 473
Db 400 TPTTSSPSSPPAPSAARGSTPAVLRRRRDAGNATTP---VPPTAPGKSLGTNNP 456

QY 474 TQAQVQAYDNRIIRINILEDLSKAWCREQRAALVWNEISKINPTSVMSMINRNPVSA 533
Db 457 ATVQIQFAYDSURRQINRMGLDLARAWCLEQRQONWVLRLETKINPTTVMSIIYKAVAA 516

QY 534 KRIGDIVSNICIVVDQTSVLHKSRLLSASDEKCFRBPVTFKFMNDSTIYKQQLGN 593
Db 517 KRLGDIVISQCVPNQATVTLRKSMR--VPGSETMCSRPLVSFSFINDTKTYEQQLGTD 575
```

```
QY 594 NEILLTTYLETQCOENTYFFQAKTDMYIKYNYEHLKTVPLSSITLTDFTIALNFTLLEN 653
Db 576 NEIFLTRKMTVCQATSQYFYQSGNEIHVYNDYHHFKTIELDGIATLQTFISLNTSLLEN 635

QY 654 VDFKVIELYTRDEKRLSNVFDIETMFREYNYAQRVSGRLRKDOLD--LSTNRNQFVDAFGS 712
Db 636 IDFASLELYSRDEQRASNVFDLEGITFREYNPQAQNIAGRLKDLONAVSNGRNQFVDGLGE 695

QY 713 LMDDDLGAVGQTVVNAVSGVATLFSSIVTGFNFKNPFGGMLIIVIGLVFAIFYFLTKK 772
Db 696 LMDSLGSGVQSIITNLVSTVGGLFSLVSGFISFPKNPFGGMLIIVLVAGVWILVLSLRR 755

QY 773 TKIYETAPIKMIYPEIDKLKEREGKSE---IAPISEELERIVLAMHIHQNSHMETKTR 829
Db 756 TROMSQQPVQMLYPCIDELAQHASGEGPGINPISKTELQAIMLA--LHEQNOEQOKRAAQ 813

QY 830 KDPKDSILTRAQNMLR 845
Db 814 RAAGPSVASRALQAAR 829

RESULT 11
US-10-055-364-39
; Sequence 39, Application US/10055364
; Publication No. US20020155433A1
; GENERAL INFORMATION:
; APPLICANT: Patience, Clive
; TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use
; FILE REFERENCE: 61750-379
; CURRENT APPLICATION NUMBER: US/10/055,364
; CURRENT FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: US/09/612,204
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US/60/142,736
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: US/60/168,532
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 39
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Murine herpesvirus 68
US-10-055-364-39

Query Match 39.1%; Score 1780; DB 14; Length 831;
Best Local Similarity 41.8%; Pred. No. 2.5e-143;
Matches 347; Conservative 160; Mismatches 267; Indels 56; Gaps 11;

QY 34 SPPNTATWSTESPLTGHYGTDDSHSGRGNNENRDSSEQKNKIYGSPTFPYRVCSASGV 93
Db 43 TPAQADATETPPPLUS-----TNTNRGFY-----FRVCGVAAT 75

QY 94 GDVFRFQTDHVCDDASDMVHSEGLLIYKQNIIPFMFRVKRYKVVTTSTVYNGIYSDSI 153
Db 76 GETFRFDLXTCPTQDKKHVEGLLYKINIVPIYFKIRRYKIIITQIWRGLTSSV 135

QY 154 TNQHTFYKSIPEWETEKMDTIYQCFNSRLNTGNNLLTYVDRDDINMTVFLQPDVGVTPD 213
Db 136 TKGFMATQAHEWVEGVDFDSIYQCYNSATMVVNVVRQVYVDRDGVNKTIVIRPDGLTGN 195

QY 214 VKRYGSOPELYLEPCGFWGVSRRRTTNCCLMDMFARSNPFPDFVTTATGDTVEMSPWS 273
Db 196 IQRTFSQPTIYSEPGMMPGYRVRTTVNCEIIVDMVARSMDPYNVITATGDSLELSFPQT 255

QY 274 GEDDHENKMKHEK--PWFSVINNYKVVDYQNRGTVPGLKTRIFLDRREYTLSEWKLKNS 332
Db 256 FDNTSQSTAPKRAMRVREKVKYKVDYNNRGTAPAGQSTFLETSPATYSWKTATQTA 315

QY 333 YCPITLWKAFYNGIQTEHSGSYHFVANDITASFTTSKEDMKFNTTVHCLNEEIKAEIEK 392
Db 316 TCDLVHMKTFPRAIQTAHEHSYHFVANEVATATNTPLTEVENTTSTYSCVSDQINKTISE 375
```



```
Db 633 GHRRYIFGGVYFEEYAYSHQLSRADVTTVSTFDLNTMLDEHFEFVLEVTYRHEIK 692
QY 669 LSNVFDIETMFREYNYAQRVSGRLKDLDDLSTNRNQFVD--AFGLMDDLG-AGQTVV 725
Db 693 DSGLLDYTEVQRNQLHDLRFADI-DTVIRADANAAMFAGLCAFEFGMDLGRVGVVM 751
QY 726 NAVSGVATLPSIVTGFNFINKPFGGMLIIVIGVLFAIYFLTKTKIYETAPIKMIY 785
Db 752 GVVGGVV----SAVSGVSFMSNPFGALAVGLLVAGLVAFAFPAFYVLQLRNPKALY 807
QY 786 PEIDKLKEREGKSEIAPISEE-----ELERIVLAMHIHQONSHMETKTRKDP 832
Db 808 PLTTKELKTSDPGGVGGEGEAGGDFEAKLAEMIRYMALVSAMERTHKARKKG 867
QY 833 KDSIL-TRAQNM-LRKR--SGYSNLKNAE 857
Db 868 TSALLSSKVTNMVLRKRNKARYSPLHND 896

RESULT 17
US-10-237-551-18
; Sequence 18, Application US/10237551
; Publication No. US20030165820A1
; GENERAL INFORMATION:
; APPLICANT: Day, Craig H.
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Parsons, Joseph M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF HERPES SIMPLEX VIRUS INFECTION
; FILE REFERENCE: 210121.538C3
; CURRENT APPLICATION NUMBER: US/10/237,551
; CURRENT FILING DATE: 2002-09-06
; NUMBER OF SEQ ID NOS: 254
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 904
; TYPE: PRT
; ORGANISM: Herpes simplex virus
US-10-237-551-18

Query Match 18.6%; Score 844; DB 12; Length 904;
Best Local Similarity 28.4%; Pred. No. 4.8e-63;
Matches 247; Conservative 164; Mismatches 366; Indels 92; Gaps 27;

QY 33 TSPNTATWSTESPLTGHYTHDSHGERGNENRDSEEQNKNIYGSPTFPYRVCSAG 92
Db 76 TTPPD-----ANATVAAGHATLR--AHLREIKVENADAQ-----FYVCPPT 115
QY 93 VGDVFRFQTDHVCPSAD-MVHSEGILLIYKQNIIPFMFRVVRKRVKVVTTSTV----- 144
Db 116 GATVVQFEQPRRCPTREPQNYTEGIAVVFKEKNIAPYKFKATMYKDYKTVSQVWFGHYS 175
QY 145 -YNGIYSDSTINQHTFYKSIPEWE--TEKMDTIYOCFNSLRNTGNNLLTVYVDRDINMT 201
Db 176 QFMGIFED-----RAPVFEVIDKINTKGVCRCSTAKYVRNMMETTAFAHRDDHETD 226
QY 202 VFLQVQDVGTDPDKRYGSOPELYLSPGFWGYSRRRTTNNCELMDMFAENRPPDFVFTA 261
Db 227 MELKPAK-VATRTSGWHTDLKYNPSVEAPHRYGTTVNCIVBEVDARSVYVDEFVLA 285
QY 262 TGDVTVMSPFWS-GEDDH-ENKMHKPKMFSVINNY-KVVDYQNRGTVPGLKTRIFLDRE 318
Db 286 TGDVFMSPFYREGSHTESYAADRFKQVDGYARDLTITKARATSP--TTRNLLTTP 343
QY 319 EYTLSEWHLKMSYCPILTKAFYNGIQTEHSGSYHYFVANDITASFTT-----SKED 371
Db 344 KFTVAMDVWPKRPVACTMTKMOEVDMLRAEYGGSGFRSSDAISTFTTNTLTEXSLRVD 403
QY 372 MKEFNTVHCLNEEBKAEITEKYYA-KVNSTHSGYGLDYFKYKTDGGLYLWQPLONRLLD 430
Db 404 LGD-----CIGRDAREIDRFARKYNATHIKVGQPOQYIATGCGFLIAYOPLLSNTLAE 457
QY 431 AKNKLNNETYSR-RSRROAESTTDPMMEM-TGNGAGGEYSSENSITVAQVQVAYDNLIR 488
```

```
Db 458 ---LYVREYMRQDRKPRNATPAFLREAPASANASVERIKTTSSTIEFARLQFTYNHQRH 513
QY 489 INNTELDLSKWCAREQHPAALVWNELSKINTSVMSMIYNRPVSAKRIGDVISVNCIYV 548
Db 514 VNDMLGR:IAVAMCELQNHLETLNNEARKLNPNATASATVGRVRVSARMILGDMNAVSTCPV 573
QY 549 DQTSVSLHSLRLLSASDEKCFSPPTVTFKFMNDSTIYKGLGVVNEILLTTTYLETQOE 608
Db 574 APDNVIVQNSNR-VSSRPCTCYSRPLVSFRVEDQGLIEGQCGENNEIRLTRDALEPCTV 632
QY 609 NTEYFQAKTDMYIYKNYEHKLTVPSSITLDTFIALNFTLLENVDFKVIELYTRDEKR 668
Db 633 GHRRYIFGGVYFEEYAYSHQLSRADVTTVSTFDLNTMLDEHFEFVLEVTYRHEIK 692
QY 669 LSNVFDIETMFREYNYAQRVSGRLKDLDDLSTNRNQFVD--AFGLMDDLG-AGQTVV 725
Db 693 DSGLLDYTEVQRNQLHDLRFADI-DTVIRADANAAMFAGLCAFEFGMDLGRVGVVM 751
QY 726 NAVSGVATLPSIVTGFNFINKPFGGMLIIVIGVLFAIYFLTKTKIYETAPIKMIY 785
Db 752 GVVGGVV----SAVSGVSFMSNPFGALAVGLLVAGLVAFAFPAFYVLQLRNPKALY 807
QY 786 PEIDKLKEREGKSEIAPISEE-----ELERIVLAMHIHQONSHMETKTRKDP 832
Db 808 PLTTKELKTSDPGGVGGEGEAGGDFEAKLAEMIRYMALVSAMERTHKARKKG 867
QY 833 KDSIL-TRAQNM-LRKR--SGYSNLKNAE 857
Db 868 TSALLSSKVTNMVLRKRNKARYSPLHND 896

RESULT 18
US-10-121-988-18
; Sequence 18, Application US/10121988
; Publication No. US20030068327A1
; GENERAL INFORMATION:
; APPLICANT: Hosken, Nancy Ann
; APPLICANT: McGowan, Patrick
; APPLICANT: Sleath, Paul R.
; APPLICANT: Mossman, Sally P.
; APPLICANT: Evans, Lawrence S.
; APPLICANT: Swanson, Ryan M.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF HERPES SIMPLEX VIRUS INFECTION
; FILE REFERENCE: 210121.538C1
; CURRENT APPLICATION NUMBER: US/10/121,988
; CURRENT FILING DATE: 2002-04-11
; NUMBER OF SEQ ID NOS: 183
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 904
; TYPE: PRT
; ORGANISM: Herpes simplex virus
US-10-121-988-18

Query Match 18.6%; Score 844; DB 15; Length 904;
Best Local Similarity 28.4%; Pred. No. 4.8e-63;
Matches 247; Conservative 164; Mismatches 366; Indels 92; Gaps 27;

QY 33 TSPNTATWSTESPLTGHYTHDSHGERGNENRDSEEQNKNIYGSPTFPYRVCSAG 92
Db 76 TTPPD-----ANATVAAGHATLR--AHLREIKVENADAQ-----FYVCPPT 115
QY 93 VGDVFRFQTDHVCPSAD-MVHSEGILLIYKQNIIPFMFRVVRKRVKVVTTSTV----- 144
Db 116 GATVVQFEQPRRCPTREPQNYTEGIAVVFKEKNIAPYKFKATMYKDYKTVSQVWFGHYS 175
QY 145 -YNGIYSDSTINQHTFYKSIPEWE--TEKMDTIYOCFNSLRNTGNNLLTVYVDRDINMT 201
Db 176 QFMGIFED-----RAPVFEVIDKINTKGVCRCSTAKYVRNMMETTAFAHRDDHETD 226
```

Qy	202	VFLQPVGDVTPDKVYIGSQOPELYLPEGWFNGSYRRRTTVNCELMDMPARSNPDPDFVTA	261
Db	227	MELKPAK-VATRTSGWHTDLDKYNPSVEAFHRYGTIVNCIVEEDVARSVYDFEVLIA	285
Qy	262	TGDTVMSPFWS-GEDDH-ENKMEKMPWFVSVINNY-KVVDYQNRGTVPVLGKTRIFLDRE	318
Db	286	TGDFVYMSFPYGRSGSHTSIAADRFQVDGFYARDLTTKARATSP--TTRNLLTTP	343
Qy	319	EYTLSEWEHLKNWSCPJLTLKAFYNGIQTEHSGSYHFVANDITAGFTT-----SKED	371
Db	344	KFTVAMDWPVKRPAYCTMTKQVEDMLRAEYGGSPFSSDAISTFTTTLNTEYSLSRD	403
Qy	372	MKEFNFTYHCLNEBJKAELEKYYA-KVNSTHYSKYGDLKPKYKTGGGLYVWLOPIQNRLLD	430
Db	404	LGD-----CIGRDAREADRNFARKYNATHIKVQPYLATGCGFLIAYQJLSLTLAE	457
Qy	431	AKNKLNNETYSR-RSRROAESTTDDPMEM-TGNGAGGEYSSENSITVAQYQVAYDNLNR	488
Db	458	---LYVREYMEQORKPRNATPAPLREAPSANASVERIKTSSIEFARLQFTYNYHIQRH	513
Qy	489	INNILEDLSKAWCRQHRALVWNELSKINPTSVMSMIYNRPYSAKRIGDVISVSNCIUV	548
Db	514	VNDMLGRITAVACELQNHHELTLWNEARKLNPAITASATVGRVSRARMLGDVMAVSTCPV	573
Qy	549	DOTSVSLHSLRLLSASDEKFSRPPVTFKPMNDSTIYKQQLGVNNEILITTTTYLETQCE	608
Db	574	APDNVIVQNSMR-VSSRPCTCVSRPLVSFYEDQGPLEGOLGENNELRLTRDALEPCTV	632
Qy	609	NTEYFQAKTDMYIYKNYHEHKTVPLSITTLDTFIALNFTLLENVDVKVIELYTRDEKR	668
Db	633	GHRFYIFGQGYVYFEYAYSQHSRADVTTFIDLNTIMLEDHEFVPLEVYVTRHEIK	692
Qy	669	LSNVFDIETMFREYNYAQRVSGLRKDLLDLSLTKRNOFVD--AFGSLMDDLGA-VAGQTV	725
Db	693	DSGLLDYEVQRNQLHDLRPADI-DTVIRADANAAMPAGLCAFFEGMGDLGRAVGKVM	751
Qy	726	NAVSGVATLFSISVTGFNFINKNPPGGMLMIIVVIGVLFAIYFLTKTKTIYETAPIKMTY	785
Db	752	GVVGGV---SAVSGVSSFSMNPFGALAVGLLVLAGLVAAPFAFRYVLQORNPMKALY	807
Qy	786	PEIDKLKREGKSEIAPISSE-----ELERIVLANWHIHOONSHMETKTRDP	832
Db	808	PLTTTELKTSDPGVGGEAGEAGGGFDEKLAAREMTRYMALVSAMERTEHKARKG	867
Qy	833	KDSIL-TQAQNM-LRKR--SGYSNLKNAE	857
Db	868	TSALLSSKVTNMVLKRNKARYSLPHNED	896

RESULT 19

```

RESD11 13
US-09-147-052-4
; Sequence 4, Application US/09147052
; Patent No. US20010014335A1
; GENERAL INFORMATION:
; APPLICANT: SAITOH, Shuji
; APPLICANT: TSUZAKI, Yoshinari
; APPLICANT: YANAGIDA, No. US20010014335A1orU
; TITLE OF INVENTION: NOVEL FUSED PROTEIN, GENE THEREFOR, RECOMBINANT VECTOR
; TITLE OF INVENTION: NOVEL FUSED PROTEIN, GENE THEREFOR, RECOMBINANT VECTOR
; FILE REFERENCE: 981167
; CURRENT APPLICATION NUMBER: US/09/147.052
; CURRENT FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: JP 08-103548
; PRIOR FILING DATE: 1996-03-29
; PRIOR APPLICATION NUMBER: PCT/JP97/01084
; PRIOR FILING DATE: 1997-03-28
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1086
; TYPE: PRT
; ORGANISM: hybrid
US-09-147-053-4

```

	Query Match.	16.6%; Score 754; DB 9; Length 1086;
	Best Local Similarity	25.9%; Pred. No. 3.3e-55;
	Matches	Conservative 156; Mismatches 375; Indels 114; Gaps 22;
Qy	51 YGTHDSHGGRNNENRDEEQNKNIYGSPTFPYRVCSASGVGDVFRFOTDHVCDDASD	110 : : : : : : : : : : : : : : : : : : : :
Dd	17 YGNSSP-----STQNVTSEVVSSVOLSEESTFYLCPPPGVGSTVIRLEPPPKCEPRK	71 : : : : : STQNVTSEVVSSVOLSEESTFYLCPPPGVGSTVIRLEPPPKCEPRK
Qy	111 MVH-SEGILLIYKQNIIPFMFRVKRYKVTTTSTVANGIYSDSITNOHTPYKSIEPEW-T	168 : : : : : : : : : : : : : : : : : : : :
Dd	72 ATEWGEGIALFKENISPYKFVLYLYKNLIQTITWTGYTRQITNRYTDRTVPUSIEEIT	131 : : : : : : : : : : : : : : : : : : : :
Qy	169 EKMDTTYQCNSURLNTGNNLLTYVDRDDINMTVFLOPDGVTPDWKRVYSQSOPELYLEPG	228 : : : : : : : : : : : : : : : : : : : :
Dd	132 DLIDGRCRSSKARYLRNVVYFAFDORAGEKQVLKPSKFNFTPESRAWHITTNETYT---	188 : : : : : : : : : : : : : : : : : : : :
Qy	229 WFWGS---YRRRTTVNCELMDFEARSNPPDPFVTATGDTVENSPM--SGEDDHENKWH	293 : : : : : : : : : : : : : : : : : : : :
Dd	189 -WVGSPWIYRTGTSVNCIVEEMDARSVPFYSYPAMANGDIANTSPYGLSPPEAAAEPMG	247 : : : : : : : : : : : : : : : : : : : :
Qy	284 EKPWFVSVINNYKVVDYQNRGTVPGLKTRIFLDREBYETLSWEHLKNMSYCPLTLWKAFY	343 : : : : : : : : : : : : : : : : : : : :
Dd	248 YPODNFKQLDSYFMDLDRKRASLPVKRNFLITSHTFGWDWAPKTRVCSMTKKKEVT	307 : : : : : : : : : : : : : : : : : : : :
Qy	344 NGIQTHSGSYHFVANDITASFTTSKEDMKEFNTTY-----HCLNEEKAEITEKKY	394 : : : : : : : : : : : : : : : : : : : :
Dd	308 EMRLATVNGRYRFMARELSATFIS-----NTTFEPNRIILGOCIKREAAAIEQIF	359 : : : : : : : : : : : : : : : : : : : :
Qy	395 -AKVNSTHSKYGDLYKFKTDGGLYLVWOPLIQ-----NRLLDAKNKLN	436 : : : : : : : : : : : : : : : : : : : :
Dd	360 RTKYNDSHVKVGHQVYLALGGPIVAYQVLSKLAHMYLRMLRDNRDTEMULDLYNNKH	419 : : : : : : : : : : : : : : : : : : : :
Qy	437 -----NETYSRSRRQAESTTDPMMEMTGNGAGGEYSSENSITVAOVQVAYONLRIINN	491 : : : : : : : : : : : : : : : : : : : :
Dd	420 AIYKKNATSLSRLRRDIRNAPNRKITLDDTTA---IKSTSSVQFAMLQFLYDHIQTHIND	476 : : : : : : : : : : : : : : : : : : : :
Qy	492 ILBDLSKACREOHRALVYNELSKINPTSVMGMIYNRPVSAKRIGDVISVNCIVVDQT	551 : : : : : : : : : : : : : : : : : : : :
Dd	477 MFSRIATAWCELQRNELVLVHEGIKINPSATASATLGRRAAKMLGDVAAVSCTAIDAE	536 : : : : : : : : : : : : : : : : : : : :
Qy	552 SVSLHSLALLSADSEKFCSPRPVTEKFWNDSTIYKGOLGVNNEILLTTTYLETCCOENTE	611 : : : : : : : : : : : : : : : : : : : :
Dd	537 SVTLQSMRMIT-STNTCYSRPLVLFSGENQGINQGQGENNELLPTEAVEPCSANHR	595 : : : : : : : : : : : : : : : : : : : :
Qy	612 YYFOAKTDMYIYKNYBEHLXTPVLSSTITLDTFIALTFTLLENDFKVIETYRDEXRLSN	671 : : : : : : : : : : : : : : : : : : : :
Dd	596 RYPLFGSGYALPENYNFVKWDAADIQIASTFVELNLTLLEOREILPLSVYTKHEELRDVG	655 : : : : : : : : : : : : : : : : : : : :
Qy	672 VFDIETMFREYNYAQRVSGRLKDLDDLSTNRNQFVDAFGSLMDDLGAQGVQTVVNAVSGV	731 : : : : : : : : : : : : : : : : : : : :
Dd	656 VLDYAEBVARRNQLHELKDYDINK-VIEVDTRY-----AGLBFGCMSITKKDA----	702 : : : : : : : : : : : : : : : : : : : :
Qy	732 ATLFSSIVTGFINFINKNPFGG-----MLMTIIVIGVLPALFYFLTCKTIYETAPIKM	783 : : : : : : : : : : : : : : : : : : : :
Dd	703 -----NPNNQOTQEAREMEITLDLINAKAMTLASLDQYAKI--EASLSS	744 : : : : : : : : : : : : : : : : : : : :
Qy	784 IYPEIDKLKEREGCKSIAPISEEELERIVLAMHHHQONSHMETKTRKQP-----KD	834 : : : : : : : : : : : : : : : : : : : :
Dd	745 AYSEAETVNNLN----ATLEQLKVAKTNLLESAINQANTDKTTFDNEHENLVEAYKALKT	800 : : : : : : : : : : : : : : : : : : : :
Qy	835 SILTRAQNMRLKRS-GYSNLKNAESVEMULN	863 : : : : : : : : : : : : : : : : : : : :
Dd	801 TLEORATNLEGLSSTAYNQIRN-NLWDLYN	829 : : : : : : : : : : : : : : : : : : : :

RESULT 20

RECD: 120
 US-10-055-364-48
 ; Sequence 48, Application US/10055364
 ; Publication No. US2002015543A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Patience, Clive
 ; TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use


```
; FILE REFERENCE: 61750-379
; CURRENT APPLICATION NUMBER: US/10/055,364
; CURRENT FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: US/09/612,204
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US/60/142,736
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: US/60/168,532
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 48
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Suid herpesvirus 1 - bases 491-850
US-10-055-364-48

Query Match
Best Local Similarity 9.7%; Score 442; DB 14; Length 359;
Matches 108; Conservative 83; Mismatches 152; Indels 24; Gaps 8;

QY 441 SRSRQ--AESTTPMTEMGTGAGGEYSSENSITVAQVAYDNLRIRINNILEDSK 498
Db :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
6 ARRARRSPGAGTPEP---PAVNGTGHRLITTCGSAEFARLQFTYDHIQAHVNDMLGRIAA 62

QY 499 ANCREQHRAALVWNELSKINPTSVMSIYNNRPVSAGKIGDVISVNCIIVDQTSVSLHKS 558
Db :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
63 ANCELQNKDRTUWSEMRSLNPSAVATAALGQRVARMGLGDVMAISRCVEV-RGGVYVQNS 121

QY 559 LRLSASDEKCFRRPVPVTFKPMNDSTIYKQGLGVNNEILLTTTLETCTOENTYVFOAKT 618
Db :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
122 MR-VPERGTCVSRPLVTEH-NGTGVIEGQGDNDNELLISRDLEPCGTGNHRRYFKLGS 179

QY 619 DMVYKYNHEHKTVPVLSSTITLDTFTALNFTLLENVDFKVBELYTRDEKRLSNVFDIETM 678
Db :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
180 GYVYEDYNNYRMVEVPE--TISTRTVLTNLTLEDREFLEPLSVYTRTELADTGLLDYSEI 237

QY 679 FREYNYA-----ORVSLGRKDLDDLSTNRNQFVDAFGLSMDDDLGAVCQTVNAVSGVA 732
Db :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
238 QRNRQLHAKFVDIDRVKVDNVLRLGIANFF-----QGLGDVGAAGKVVLGAT 289

QY 733 TLFSSIVTGFINFKNPFGLMMLIIVIGVLFALVFLTKTKIYETAPIKMIYPEIDKLG 792
Db :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
290 GAVISAVGGMVSFLSNPFGALGULLVLAGLVAFLAYEHSRLRRNPKALYPTVTKTL 349

QY 793 EREGKSE 799
Db :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
350 KEDGVDE 356

RESULT 21
US-10-127-733-2
; Sequence 2, Application US/10127733
; Publication No. US20020150889A1
; GENERAL INFORMATION:
; APPLICANT: EICKMANN, Markus
; APPLICANT: GICKLHORN, Dorothee
; APPLICANT: RADSAK, Klaus
; APPLICANT: HAUSER, Hans-Peter
; APPLICANT: GIESENDORF, Bernhard
; TITLE OF INVENTION: IMMUNOREACTIVE REGIONS OF GLYCOPROTEIN GP11 OF
; TITLE OF INVENTION: VARICELLA ZOSTER VIRUS (VZV)
; FILE REFERENCE: 058315/0128
; CURRENT APPLICATION NUMBER: US/10/127,733
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US/09/219,797
; PRIOR FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 206

; FILE REFERENCE: 61750-379
; CURRENT APPLICATION NUMBER: US/10/055,364
; CURRENT FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: US/09/612,204
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US/60/142,736
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: US/60/168,532
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 48
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Suid herpesvirus 1 - bases 491-850
US-10-055-364-48

Query Match
Best Local Similarity 9.7%; Score 350; DB 14; Length 206;
Matches 68; Conservative 55; Mismatches 75; Indels 2; Gaps 2;

QY 476 AQOVAYDNLRIRINNILEDSKAWCREQHRAALVWNELSKINPTSVMSIYNNRPVSAGR 535
Db :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
3 AMLQFTYDHIQEHVNEMLARISSSWCQLQNERALWSGLFFINPSALASTILDQRVKARI 62

QY 536 IGDVISVNCIIV-DQTSVSLHKSRLLSASDEKCFRRPVPVTFKPMNDSTIYKQGLGVN 594
Db :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
63 LGDVIFVSNCPGLGSDTRIILQNSMR-VSGSTTRCYSRPLISIVSLNGSGTVEGQLGTDN 121

QY 595 EILLTTTLETCTOENTYVFOAKTDMVYKYNHEHKTVPVLSSTITLDTFTALNFTLLENV 654
Db :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
122 ELIMSRDLLEPCVANHKRYFLFGHHYVYEDYRVYREIAVHDVGMISTYVDLNLTLKDR 181

QY 655 DPKVIELYTRDEKRLSNVFD 674
Db :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
182 EFMLQVYTRDELDTGLLD 201

RESULT 22
US-09-815-242-13696
; Sequence 13696, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13696
; LENGTH: 526
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13696

Query Match
Best Local Similarity 2.8%; Score 127; DB 9; Length 526;
Matches 111; Conservative 104; Mismatches 228; Indels 174; Gaps 29;

QY 270 PFWGEGDDHKNKME--KPMFVSVNNYKVVYDQNRGTVPGLGKTRIF-----LDREYTLS 323
Db :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:| :|||:|
6 PAWYGS---ERTWHADITPWFSHFR---LEFDD----TPHQIRLFOEQIDSLRLVLA 54
```



```
; SEQ ID NO 2
; LENGTH: 3907
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-171-311-2

Query Match      2.8%; Score 126.5; DB 15; Length 3907;
Best Local Similarity 18.7%; Pred. No. 1.5;
Matches 110; Conservative 101; Mismatches 196; Indels 181; Gaps 26;

QY 373 KEFTTYHCLNEIEKAEIEKKYAKVNSTHSGYDLYKPKTDGGLVYWQPLIQNRLLDAAK 432
Db 605 KKNVLDMAESQEAELERLTQLFSHEE--ELSKLKED-----LEIE 647
QY 433 NKLNNETYSRRSRQAESTTDPMMEMTNGAGGEY-----SSENSITVAQVQVAYDNL 485
Db 648 HRIN-----IEKLDNLGIHYKQIDGLQNMESQKIETMQEKNL 688
QY 486 RIRINNILEDLSKAWCREQ-----HRAALVNMELSK-----INPT 520
Db 689 ITKQNLILEISKLQDLQSLVNSKSEEMTLQINELQKEIILRQEEKEKGTLEQVQEL 748
QY 521 SVMSMYNRPVSAKRIGDV-----ISVSNCLVVDQ-----TSVSLHKSRLLS 563
Db 749 QUKTELLEKQMKKE--NDLQEKFAQLEAENSILKDEKKTLEDMLKIHTPVSQEERLIFLD 807
QY 564 ASDEKCFSPVPVTFK----FMNDSTIYKGO--LGVNNEILLTTTLYLETQENTYVYFQ--- 615
Db 808 SIKSK--SKDSVWEKEIEILLIENEDLKQCCIQLNEEIEKORNTFSPAENKFEVNYQELQ 865
QY 616 -----AKTDMYIYKNYEHK-----TVPL--SSITLTD--TFIA 645
Db 866 EYVACLLKVKDDLEDKQKQLEYSKALKALNEELHLQINPTTVKMKSSVDFDEKTFVA 925
QY 646 LNFTLLENVDKVIEL-----YTRDEK-----RLSNVFD--IETMFREYNYAQRVSGLRK 694
Db 926 ETLEMGVEVEKDTTLEMEKLEVTREKLELSQRLSDLSQKQKGEISFLNEEYKSLQ 985
QY 695 DL-----LDLSTNRNQFVDAFGSLMDLDGAVGQTVVNAVSGVATLFSIVTGFINF 745
Db 986 EKEQVSLRCRELEIINHNR-----AENVQSCDTQVSSLLDGVVMTSGAGSGVSK 1037
QY 746 IKNPFGGMLMIIV-----VIG-----VLFAIYFLTKTKIYETAPIK--MIYP 786
Db 1098 ELNVLKSEQNDLRL-----QMEAQRICLSLVYSTHVDQVREYMEKDK 1141
```

RESULT 25

```
US-10-171-311-8
; Sequence 8, Application US/10171311
; Publication NO. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glatt, Karen
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoersh, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; FILE REFERENCE: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10171,311
; PRIOR FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
```

```
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 3917
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-171-311-8

Query Match      2.8%; Score 126.5; DB 15; Length 3917;
Best Local Similarity 18.7%; Pred. No. 1.5;
Matches 110; Conservative 101; Mismatches 196; Indels 181; Gaps 26;

QY 373 KEFTTYHCLNEIEKAEIEKKYAKVNSTHSGYDLYKPKTDGGLVYWQPLIQNRLLDAAK 432
Db 605 KKNVLDMAESQEAELERLTQLFSHEE--ELSKLKED-----LEIE 647
QY 433 NKLNNETYSRRSRQAESTTDPMMEMTNGAGGEY-----SSENSITVAQVQVAYDNL 485
Db 648 HRIN-----IEKLDNLGIHYKQIDGLQNMESQKIETMQEKNL 688
QY 486 RIRINNILEDLSKAWCREQ-----HRAALVNMELSK-----INPT 520
Db 689 ITKQNLILEISKLQDLQSLVNSKSEEMTLQINELQKEIILRQEEKEKGTLEQVQEL 748
QY 521 SVMSMYNRPVSAKRIGDV-----ISVSNCLVVDQ-----TSVSLHKSRLLS 563
Db 749 QUKTELLEKQMKKE--NDLQEKFAQLEAENSILKDEKKTLEDMLKIHTPVSQEERLIFLD 807
QY 564 ASDEKCFSPVPVTFK----FMNDSTIYKGO--LGVNNEILLTTTLYLETQENTYVYFQ--- 615
Db 808 SIKSK--SKDSVWEKEIEILLIENEDLKQCCIQLNEEIEKORNTFSPAENKFEVNYQELQ 865
QY 616 -----AKTDMYIYKNYEHK-----TVPL--SSITLTD--TFIA 645
Db 866 EYVACLLKVKDDLEDKQKQLEYSKALKALNEELHLQINPTTVKMKSSVDFDEKTFVA 925
QY 646 LNFTLLENVDKVIEL-----YTRDEK-----RLSNVFD--IETMFREYNYAQRVSGLRK 694
Db 926 ETLEMGVEVEKDTTLEMEKLEVTREKLELSQRLSDLSQKQKGEISFLNEEYKSLQ 985
QY 695 DL-----LDLSTNRNQFVDAFGSLMDLDGAVGQTVVNAVSGVATLFSIVTGFINF 745
Db 986 EKEQVSLRCRELEIINHNR-----AENVQSCDTQVSSLLDGVVMTSGAGSGVSK 1037
QY 746 IKNPFGGMLMIIV-----VIG-----VLFAIYFLTKTKIYETAPIK--MIYP 786
Db 1038 VNKSGEESKIMVEDKVSFENMTVGBESKQEQILDLHPLSVTKESSLRATQPSENDKLOK 1097
QY 787 BIDKLKEREKGEIAPISEEELERIVLAM-----HIHQONSHMETKTRK 830
Db 1098 ELNVLKSEQNDLRL-----QMEAQRICLSLVYSTHVDQVREYMEKDK 1141
```

RESULT 26

```
US-10-171-311-6
; Sequence 6, Application US/10171311
; Publication NO. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glatt, Karen
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoersh, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; FILE REFERENCE: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035
```

Query Match	2.7%	Score 125;	DB 9;	Length 1009;
Best Local Similarity	17.3%	Pred. No. 0.24;		
Matches 114;	Conservative 112;	Mismatches 231;	Indels 202;	Gaps 30;
Qy	267	EMSPFWGEGDDHENKMKHEKFWFSVINNYKV-VDYQNRGTVPGLKTRIFLDRREYTL	SW 325	
Db	459	ELDKLLSGKOELENOINEEK---TILKNLEIKLDYDKSKLDLN-----DKSFISEIK	509	
Qy	326	KHLKNMSYCPLTWLKAFYNGIOTHEGSGSYHFVANDITASFTTSKEDMKFNTTVHCL---	382	
Db	510	SAVKIGIQCPIC-----GNEIQDLG---HHIDFD---SIKRQNEIKEIANIHTMESN	557	
Qy	383	---NEEIKAEIEKKYAKVNSTHSGYGDLYKFTDGLYLVMQPLIQNRLLDANKLNNE	438	
Db	558	IAVNSEIKF-VNEIKSNIN-----IKTQDLSL---EVLNKRLLLENALNNQ	602	
Qy	439	---TYSRRRRQAESTT---DPMEMTNGAG-----GEYSSENSITVA	476	
Db	603	RLNKFIEQKMEKDNLTLOIHKNQLRLNKSELSKICRLNLITEFETLSKYNNITNEVD	662	
Qy	477	QVOYAYD-NLRIRINILED---LSKAWCRBOHRAALVWNELSKINPTSVMSMYNRPV	531	
Db	663	YKKYIODVNOHQHBSNQIEDKLIQLSORKLIEQNNLNHYENQLE-----TYNNDL	712	
Qy	532	SAKRIGDIVISNCIVVDQTSV-----SLH	556	
Db	713	ELNEQSIEMEMSRNLNTDDNINEITAWRGEQBELEQKRDYKKRYHEFEMEIALESLT	772	
Qy	557	KSURLLSASDEKCFSRPPVTFKPMNDSTIYKGQGV-----NN-----BILL	598	
Db	773	KDKELLDSD-----KLKDDYELKKGKMWTLIDEYSVAHYQCNNIKKTQSIVS	820	
Qy	599	TTTTYLETQCENTEYFQAQKDMWIKYKNEYHLKTVPLSSITTLDTFALNFTLLENVDFKV	658	
Db	821	HINYLNQELKDDQEI FOL-AEILSGKNKNKL-----TLE-----NEVLIIYLD---	862	
Qy	659	IELYTRDKELSNVFIETWTFREYNYAQRVSGLRKDLDLSTNRNQFVDAFGLSMDLGL	718	
Db	863	-QIIAQANLRLATMSDNRYQLIRREAVSHGLSGLEIDVDFLHNSKSH-----	909	
Qy	719	AVGQTVNAVSGVATLFSS---IVTGFINFKPFGMLMIIVIGULFAIFYLTKTKTIY	776	
Db	910	-----ISSLSGETTQSSIALAALGSEIVQOOSGGISLSEIFIDEGF---TLDOETL	959	
Qy	777	ETAPIKMIYPEIDKLKEREKGSB---IAPISBEELERIVLAWHIHOONSHMETKTKD	831	

[illegible]

Db 960 ETV-----LDTLLNLKSTGRMVGIISHVSELK-NRIPLVLEKSDQYQSSTFRKN 1009

RESULT 28

US-09-815-242-5251

; Sequence 5251, Application US/09815242

; Patent No. US20020061569A1

GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Karl L.

; APPLICANT: Zyskind, Judith W.

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John D.

; APPLICANT: Cart, Grant J.

; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes in

; FILE REFERENCE: ELITRA.011a

; CURRENT APPLICATION NUMBER: US/09/815,242

; PRIOR FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 5251

; LENGTH: 996

; TYPE: PRT

; ORGANISM: Staphylococcus aureus

US-09-815-242-5251

Query Match 2.7%; Score 124; DB 9; Length 996;

Best Local Similarity 17.4%; Pred. No. 0.28;

Matches 109; Conservative 106; Mismatches 216; Indels 196; Gaps 28;

QY 267 EMSPEWSGEDDHENKMKPWFVSVINNVKV-VDYQNRGTVPGLKTRIFLDREEVTLSE 325

Db 459 ELDKLLGSQLEENINEEK---TILKLEIKLDYDKSKLDN-----DKESFISEIK 509

QY 326 KHLKMSVCPLFLWKAIFYNGIQTEHSGSVHFVANDITASFTTSKEDMKFNTTYHCL--- 382

Db 510 SAVKIGDQPCIC-----GNEIQDLG--HHIDFD---SIKRQNEIKEIEANHTMESN 557

QY 383 ----NEEIKAEIKKYAKVNSTHSKYGLKYPKTDGGLYLVWQPLIQNRLLDAKKNLNE 438

Db 558 IAVHNSKTF-VNEKISIN-----IKTQSDLSL-----EVLNKRLLLENALNNQ 602

QY 439 ----TYSRRSRQABSTT-----DPNMEMTGNAG-----GEYSSENSITVA 476

Db 603 RLNKFIEQMKSEKNLTQIHNKQLRLNKNSELIKICNLITEFETLSKYNNTNFEVD 662

QY 477 QVOYAYD-NLRIRINILED-----LSKAWCREQHRAALVWNLSKINPTSVMSMYNRPV 531

Db 663 YKKYIQDVNQHQHNSQIEDKLIQSQRKLIENNNLHYENQLE-----TYNNDL 712

QY 532 SAKRGDVLVSNVICVVDTSV-----SIH 556

Db 713 ELNEQSIEMEMRLNLTDDNDINEIAMWGEQEELEQKRDYKRYHFEFEMEIALESIT 772

QY 557 KSLRLLSASDEKCFSPVPVTFKFMNDSTIYKQLGV-----NN-----ETLL 598

Db 773 KDKELLDSD-----KLKDDYELKKGKNTLIDEVSAVHYQCONNIKKTQSIVS 820

QY 599 TTTYLETQCENTEYFQAKTDMYIKNYEHLKTVPLSSITTLDTFFIALNFTLLENVDFKV 658

Db 821 HINYLQELKQOEIFOL-ABILSGKNKNL-----TLE-----NFVLIYLD--- 862

QY 659 IELVTRDEKRLSNVDFDIETMFREYNYAQRVSGRKDLDLSTNRNOFVDAFGSLMDDLG 718

Db 863 -QIIAQANLRLATMSDNRVQLIRREAVSHGLSGLEIDVFDLHNSKSH----- 909

QY 719 AVGQTVNAVSGVATLFSS--IVTGFINFKNPFPGMLMIIVIGVLFAIYFLTKT--- 773

Db 910 -----ISSLSGGETFOSSIALALGLSEIVQQQSGGISLESIFIDEFGT--LQOETLET 961

QY 774 -----KIYETAPIKMIYPEIDKLER 794

Db 962 VLDTLNLKSTGRMVGIISHVSELKNR 988

RESULT 29

US-10-242-056-59

; Sequence 59, Application US/10242056

; Publication No. US20030113323A1

GENERAL INFORMATION:

; APPLICANT: Ensign, Jerald C

; APPLICANT: Bowen, David J

; APPLICANT: Petell, James

; APPLICANT: Fatig, Raymond

; APPLICANT: Schoonover, Sue

; APPLICANT: French-Constant, Richard

; APPLICANT: Orr, Gregory L

; APPLICANT: Merlo, Donald J

; APPLICANT: Roberts, Jean L

; APPLICANT: Rocheleau, Thomas A

; TITLE OF INVENTION: Insecticidal Protein Toxins from

; NUMBER OF SEQUENCES: 88

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: DowElanco

; STREET: 9330 Zionsville Road

; CITY: Indianapolis

; STATE: IN

; COUNTRY: US

; ZIP: 46268

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; FILING DATE: US/10/242,056

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/063,615

; FILING DATE: 18-MAY-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/395,497

; FILING DATE: 28-FEB-1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/007,255

; FILING DATE: 06-NOV-1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/608,423

; FILING DATE: 28-FEB-1996

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/705,484

; FILING DATE: 28-AUG-1996

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/743,699

; FILING DATE: 06-NOV-1996

; ATTORNEY/AGENT INFORMATION:

	Query Match	Best Local Similarity	2.7%; Score 123; DB 9; Length 792;	
	Matches 108; Conservative 93; Mismatches 250; Indels 144; Gaps 17;			
Qy	450	----	STTDPMM-----MTNGAGGEY-----SSNSTTVA	476
Db	764	SLLDYTLQADPSLEADLVTDGKSEPMDFNGSLYFELFPLFPLVATRFANEQQFSPA	823	
Qy	477	Q--VOYAYDNLRIINNILEDLSKAWCREHQRAALVWN-----ELSK-----INPT	520	
Db	824	QKSLHYIFD-----PAMKNKPHNAPAYWVRPLVEGNSDLSRHLDDSIDPD	869	
Qy	521	SVMSMIYNRPVSAKRGDVISVNCIV-----VDQTSVSLHKSRLRLLSASDE	567	
Db	870	---TQYAHFVYQKAVFIAYVSNLTIAOGDMWYRQLTRDGLTQARVYVYNLAELLG---	922	
Qy	568	KCFSRPVPVTFKM---NDSITYKGQGV-----NNEILL-----TTTTLETQCE	608	
Db	923	---PRDVSLSGIWTFQTLDTLIAAGQKAVLRDFEHQLANSQTALPALPGRVNVTYKLADN	979	
Qy	609	NTEYFFQAKTDMYIKYNEHLK-----	630	
Db	980	G---YFNEPLNMLSHWDTLLDARLYNRHNLVTDGKPLSLPLYAAPVDPVALLAORAOS	1036	


```

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2175
; TYPE: PR1
; ORGANISM: Mus sp.
; FEATURE:
; OTHER INFORMATION: CLASP-1
US-09-978-244A-4

Query Match      2.6%; Score 119; DB 11; Length 2175;
Best Local Similarity 18.6%; Pred. No. 2.6;
Matches 173; Conservative 129; Mismatches 351; Indels 276; Gaps 42;

QY 24 SLSTAETGVTSPPTATWSTESPTTGHVTHDSSHGGRGN--NENRDEEQNKNIYSPS 81
Db 290 SAEALAEGLDPLDNCVT--CCTLEEDSSSEHLPDAFKYLTETEDVTKTRN----- 341
QY 82 TFPYRVCSAGVGDFRFTQDHVCPDASDMVHSEGIILLYKQNIIPFMFRVKYKRVVTT 141
Db 342 -----MGRINLFLSD--PDIDTLKLRDIFENELMIKPFEEKAARIMICR 387
QY 142 STVN--GIYS-----DSITNHTYKSIETPWTETKMDTIYQCFNSLRINTGNNLLTVYDR 195
Db 388 ALNFNLOGCVTENEYDPVTNIEPFVSV-----ALYDLRDNRKISADPHV----- 432
QY 196 DDINMTVLPQVDGVTPDVK-----RYGSOPELYLEPGFPGWGSYRRRTTWNCELMD 246
Db 433 -DLNHPAVRQMLSGTPPALENGNIDTGTPOSEPHIKGLPE-BWLKPPKQA----- 482
QY 247 MFARSNPPDFPFVATGDTVMSPFWSGEDDH-----ENKMKHEK-----PWFVSVINNYK 296
Db 483 VFSVSDPHEIPLVAKVEKVLGMNIGSGAEPIYKPNPDSNKPFAQILKSNQFCSKLGK 542
QY 297 V-----VDYQNGTVPLQKTRIFIDREBYT--LSWEKHKL-----NMYSYCP 335
Db 543 MPFAWAVRSVPKQGNVDRSRFSPLRYQESSKMSSSEDLIKLVSDYRRADRISKMQSIP 602
QY 336 LTLWKAFFNGIQTTHSGSYHVFANDITASFTTSKEDMKEFNTTVHCLNEEIKAEIK--- 392
Db 603 GSLDIADVNI-IPLEH-----PNCVTSSEFI-----AQSEPTVEVEBFYI 647
QY 393 ---RYAK-----VNSTHSGYGLKYPKFTDGGLYVMQPLIQNRLLDLDAKNLNNET 439
Db 648 DSTKYCRPYRVYKQIYVYPRKLYVDSQKCF-----NKARNIT 685
QY 440 YRSRRAQAEETDPMEMTGNGAG-----GEYSSENSITVAQVQYAY 482
Db 686 VCIEFKNSDDGAKPMKCIYKPGGPLFTSSAYTAVLHHSONPDFSDEVKIELPTQLHGK 745
QY 483 DNLIRINNILLELSKAWREQHRAAL-----VNNELSKINPTSVMSIYNRPVSAKRI 536
Db 746 HHLFSPYHITCDIN-AKANAKKEALETSVGVAWLPMLMKHD--QIASQEYNIPIATLTP 802
QY 537 GDVISVNCIIVDQTSVLSHLKSLRLLSASDEKCFSRPPVTFKFMNDSTIYKQGLGVNNEI 596
Db 803 PNYLSIQ-----DPTSAC-H-----GGSIDIKWDGCKPLFKV---STFVVST--VNTQD 845
QY 597 LLTTTYLETQENTYFYFQAKTDMYIYKNYEHKLTVPSSITLDTPI--ALNFTLLENV 654
Db 846 PHVNAFFRQCKREKMSQSTSSFVRAKNLLNVDKIHMSPLPIILNQLFKILVQNE 905
QY 655 DFKVIELYTR-----DEKRLSNVFD--IETMFREYNVYQVRV--SGLRKOLLDLSTN 702
Db 906 EDEITATVTRVLADIYAKCHEEQDHSVQSYKIFVFKTKYKERTIHEELAKNLSDU--- 962
QY 703 RNQFVDAFGLMDLDAVGQTVVNAVSGVATLFSISVTGFINFTKPNPFGGMLIIVVIGV 762
Db 963 -----LKSNDSTIVKHLVHSWFF----- 981
QY 763 LFAIYFLTKTKIYETAPIKMIYPEIDKLEREKGSEIAPISBEELERIVLAWHHOONS 822
Db 982 -FAIILKSMQAHLIDTNTKIQL--PRAQRFPES-----YQSELDNLVWGL-----CD 1024

; Query Match      2.6%; Score 117.5; DB 12; Length 602;
; Best Local Similarity 16.9%; Pred. No. 0.45;
; Matches 74; Conservative 100; Mismatches 188; Indels 75; Gaps 14;

QY 354 YHEVANDITASFTTSKEDMKEFNTTVHCLNEEIKAEIKYAKVNSTHSGYGLKYPKTD 413
Db 18 YNKTNETSWTPEETL-----PVSEKQETI--TSTTTGKWEYTTD 60
QY 414 GGLYVMQPLIQNRLLDLDAKNLNNETYRSRRQAEETDPMEMTGNGAGGEYSSENSI 473
Db 61 DGKKYVYNNAITGTTWKEPNEIEBELS--NEKLAETKTELSELDKLSKPVELTGTPTSK 118
QY 474 TVAQVOYAYDNLRIRINNILLELSKAWREQHRAALVWNELSKINPTSVMSIYNRPVSA 533
Db 119 TNEARQKSF--LKLLSDN---KVNSTWSPQ---AVMENLVDPKPEYWSVKD-----PVTR 164
QY 534 KRIGDVISVS-----NCIVVDQTSVLSHLKSLRLLSASDEKCFSRPPVTFKFM---ND 582
Db 165 KQLYEYELVSKQSELSNKSLLLENFKRNFNLELRKLEAKNLVYTNTRITIKLWIDQD 224
QY 583 STIYKQQLGVNNEI--LLTTTYLETQENTYFYFQAKTDMYIYKNYEHKLTVPSSITLTD 641
Db 225 NPIFKHSMSDSGELAAIFVEYTDKLEQHEKLLQTKKQALIELSTYLRQVNSSLVEKSQ 284
QY 642 TFIALNFTLLENVDFK-----VIELYTRD--EKRLSNVFDIETMFREYNY--- 684
Db 285 TWESLYENLINDSRFOSNKNFNQNLTKNLTLQLYENEIFPRRIDDIKSKQITVISKINRND 344
QY 685 -----YAOVRVSGLR-----KDLIDLSTNRNFOVDAFGLMDLDAVGQTVVNAVSGV 731
Db 345 RKARESYKLLGELELDADTEFKDIIDKIENNDAFIEICRNGSSALELFWIDVDEKKQI 404
QY 732 ATLFSSIVGTGFINFIKN 748
Db 405 LKVKKNLVDSVISTMKN 421

RESULT 34
US-09-486-734A-4
; Sequence 4, Application US/09486734A
; Patent No. US20020164732A1
; GENERAL INFORMATION:
; APPLICANT: Chopin, Marie-Christine
; APPLICANT: Clier, Florence
; APPLICANT: Erlich, S. Dusko
; APPLICANT: Gautier, Michel
; APPLICANT: Schouler, Catherine
```


APPLICANT: Institut National de la Recherche Agronomique
; TITLE OF INVENTION: Resistance Mechanisms to Ic Type R/M
; FILE REFERENCE: Bacteriophages of Lactic Acid Bacteria
; CURRENT APPLICATION NUMBER: US/09/486,734A
; CURRENT FILING DATE: 2000-05-03
; PRIOR APPLICATION NUMBER: PCT/FR98/01873
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: FR 97/10885
; PRIOR FILING DATE: 1997-09-02
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1025
; TYPE: PRT
; ORGANISM: Lactococcus lactis
US-09-486-734A-4

Query Match 2.6%; Score 117.5; DB 10; Length 1025;
Best Local Similarity 17.5%; Pred. No. 1.1;
Matches 173; Conservative 163; Mismatches 314; Indels 339; Gaps 49;

QY 61 RGN-----ENRDEE-----QNKITYGSPFPYRVCSASGVDFRQTDHVC PD 107
DB 83 RGENGAQVLLEREDGKVTLEAFRNKDISGTS--YEV-----VHQVWD 127
QY 108 ASDMVHSEGILLI-----YKQNIIPFMRVRKRVKVTSTTVNGIYSOS- 152
DB 128 SRVRGDSVSLINGLPIIHIELQESAKDGMQAYQIRVAE-----DGFPGIYATQ 183
QY 153 ----ITNQ-HTFY-----KSIPEWETEKMDTIYQCFN-----SLRLNTG 188
DB 184 INVIPNKVDTRYFARPSDETAAYARMKFLNWRTEQDVSDFTRTVLRIPDAHE 243
QY 189 LITYV-----DRDDNMTVFLP-----VDGVTDPVKYGSQPELYLEPGFW- 235
DB 244 LISQVTLVDDPKNPKFLMALRPYQIHAIKIRKRAAQH-----EGFVHWATGSK 295
QY 236 R-----RTVNCLEMDMFARSPPEFFVTA--TGDVMSFPWSGE 275
DB 296 TITSFVATKLLAQAIGVDRTVWVDRTDLDQAQDEFTKFASEYHTGQTGNSVANTLI 355
QY 276 DDHENKMKHPFVSVINNYKV-----DYQNRGTVPGLKTR-----IPL 315
DB 356 VGKQKQQLARNLLSKNNNTILVTIQKLSAAMSAQOESEKSGNOPEKLQEHVFI 415
QY 316 DREBYTLSEKHLKQM-SYCPITLM-----KAFYNGIQTEHSGSYHFVAND-----ITASF 365
DB 416 VDEAHRAVSDEEMKRIKILPNSTWFLGTGTPIFEANKKQENGTEARTTSQQYGPLLSY 475
QY 366 TTSKEDMKE-----FNTTYHCL-----NE 384
DB 476 TT-KNAMDDGAVLGOVEYHSLISEDBLEVIYQLNKGKLPDALQOEELPAELYEKDE 534
QY 385 EIKAEIEKKY-----AKVNSTHS-----KYGDLKVFKTDGGLYLW 420
DB 535 HIRTMLOKIFNRSVVKKFKVNGPPTMSAILTTHSIAQAQHIYRLKEMKONG-----588
QY 421 QPLIONRLDANKLNNEYSR-----RERRQAESTDDPMEMTNGAGGEYSSENS 472
DB 589 -TLNLRQFDERHRLDKDFPRVAITFTSNPDRLEKNEQDDELVEIM-----KEYAKQPD 642
QY 473 ITVAQVOYADNLRIRINILEDLSKAWCREQHRAALVNE-----LSKINTSVMS 524
DB 643 ASYQDEKLYN-----QINIKRLARK--EKQYQSDGQWLDFVIVVDRLLTGFDPSPAIL 694
QY 525 MIYNRPVSAKRIGDVISVNCIWDOTS--VSLHK-----SLRLLSASDEKCP 571
DB 695 LYIDREMYQKLLQAFSRNRYITGKDSGLIVSFAPKPTMKENVQNTFFLFNEHQNFQ 754
QY 572 RPP-----VTFPMNDSTIYK--GOLGVNBEILLTTTYLETCEQENTYTYFOAKTDMYI 623
DB 755 LIPREYEEVKFEICSTLYKQSEADLSDNPHDLKTMIAQVSAYQKLGKSYKA-----F 808

QY 624 KNYEHLKTVPLSSITTLDTFIALNFTL-----LENVDFKVIETYRDEKRLSNVFD 674
DB 809 RSYDOYE-----EDFEAFSEVVEQLPOYRGCKTENVTIKEM-IEDEEHP--D 855
QY 675 IETMFREYNYAQRVSGRLKDLDLSTNRNQFVDAFGSLMDDDLGAVGTVVNAVSGVATL 734
DB 856 FEKLQEIATFSQ-LNATHKDWD-SFYINQLLKAII--QLNEAGAVEK-----899
QY 735 FSSIVTGFINIKNPPFGGMLMIIVVIGLVFAIYFLTKTKIYETAPIKMI--YPEID--K 790
DB 900 FEKEIQ-----KDP-----QIQKMYHTLKDQLVNTTEIDVAQ 933
QY 791 LKREGEKSEIAPISSEELERIVLAMHIHQ 819
DB 934 LKETSIQNEIQRLLOKAEAEFGLSFDFLQ 962

RESULT 35
US-10-087-464-10
; Sequence 10, Application US/10087464
; Publication No. US20030059436A1
; GENERAL INFORMATION:
; APPLICANT: Chishti, Athar
; APPLICANT: Oh, Steven
; APPLICANT: Liu, David
; APPLICANT: Goel, Vikas
; APPLICANT: Li, Xuerong
; TITLE OF INVENTION: Band 3 Antigenic Peptides, Malaria Polypeptides and Uses Thereof
; FILE REFERENCE: SI237/7019
; CURRENT APPLICATION NUMBER: US/10/087,464
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 06/272,930
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 1639
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-087-464-10

Query Match 2.6%; Score 117; DB 15; Length 1639;
Best Local Similarity 18.6%; Pred. No. 2.5;
Matches 148; Conservative 122; Mismatches 254; Indels 270; Gaps 39;

QY 55 DSSHGERGNN-----ENRDEEQNKNIYGSPTFPYRVCSASGVDFRQTDHVC PDAS 109
DB 912 DTSHSTNLNLSKLFPENITLSLGNKNYI-----939
QY 110 DMVHSEGLILLYKQNIIPFMRVRKRVKVTSTTVNGIYSDSITNQHTFYKSTPEWETE 169
DB 940 -----QELIQKSSSENFYEKILKDSDTF--YNESFTN--FVKS-----973
QY 170 KMDTIYQCFNSLRNTGNNLTYYVDRDINNMTVFLQVDPVDPV-KRYGSOPELYLBP 228
DB 974 KADDI-----NSLDESRRKKL-----EEDINK--LKKTLQLSFDLYNKKLRLPDK 1022
QY 229 WFWGYSR---BRTTVNCELMD--MFARSNPP-----FDFE-----VTATGDTVMSP 270
DB 1023 KTVGKYKMQIKKLTLLKEQLESKLNSLNNPKVLQNFVSFFNKKEAEIAETNLENTK 1082
QY 271 FWSGEDDHENKMKHPFVSVINNYK-VVDYQNRGTVPGLKTRIFLDREBYTLSEKHLK 329
DB 1083 I-----LLKHKGVLKYNIGESSPL-----KTLSEESIQT 1112
QY 330 NWSYCPILTKAFYNGIQTEHSGSYHFVANDITASFTTSKEDMKEFNTTYHCLNEEIKAE 389
DB 1113 EDNTVASLENFKVLSK-----LEGLKDNLNLEKKKLSYLSGLHLLIAELKEV 1160
QY 390 IEKKYAKVNSTHSGYDL-----KYFKTDDGLYLVNQPLIONRLDANKLNNEYSR 442
DB 1161 IKKNYNTGNSPSNNTDVNNALESYKFKLPEG-----TDVATVWSESGSDTLEQ 1209


```
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11522
; LENGTH: 1167
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-815-242-11522

Query Match      2.5%; Score 115.5; DB 9; Length 1167;
Best Local Similarity 17.7%; Pred. No. 1.9;
Matches 151; Conservative 118; Mismatches 267; Indels 319; Gaps 36;

QY 51 YGTHDSSHGRCNNEN-----RDSEQKNYIG-SPSTFPYRVCASGVDFRFPQ 100
Db 270 YNMTKLGDRNYENFAKKNIAARTLNERKULFGKNPELDFNLTSLR--GNIHOSI 327

QY 101 TDHVCPCDASDMVHSEGIILLYKQNIIPFMRVRYKRVKVTSTVYNGIYSDSITN- 155
Db 328 KEE---EALDWTIISH-----LIITKPIEDAIFGDNKKNPIAKA 361

QY 156 -----QHTFYKSTIEPWT 168
Db 362 LDKMWLKLSDLGETKDLKNLYESVTEAARAKSQSQBELIKNLYNTFFKEAFRKQS 421

QY 169 EKMDTIY---QCFNSLRNTGGLTYVDRD--DINMTVELQPDVGVTPDKRYGSOBEL 223
Db 422 EKLGIYTPIEVVDFILRATNGILKHHENTDFNDOSITIP-DPFWGTSGFIARLLSKEND 480

QY 224 YLEPGWFGSVRRRTTVNCELMDPFARSNPPDFVFTATGDTVEMSPWSEGDHDKNKH 283
Db 481 FIS-----DEALKEKFQGLFAFDIVLLS----- 504

QY 284 EKPWFVSVINNYKVVDYQNRGTVPGLKTRIFLDREYTLSEWKLKNMSPCLTLWKAFY 343
Db 505 ---YVIALINITQAA--QNR-----DSSLKNFKNIALT----- 532

QY 344 NGIQTEHSGSVHFVAND--ITASFTTSKEDMKFNTYTHCLNEEIKAEIKYAKVNSTH 401
Db 533 -----DSLDELEKNKDGVIPIGPFYLPEDLKE-----NKEIKTMEKQNRVVIIGN 578

QY 402 SKYGDLYFKYTDGGLYLVWQPLIQNRLLDAKNKLNNETYRSRRSQ-AESTTDPMMWTG 460
Db 579 PYPYSGAKSENNQNLNLS-HPKLEKRVY-----EYGNKSTQNKSTRTDLI----- 625

QY 461 NGAGEYSENSITVAQVOYVDNLRIRINNILEDLSKAWCREQHRRAALVWNELSKINPT 520
Db 626 -----HSIRWA-----SDLLKD-----K 638

QY 521 SVMSMIYNRPVSAKRIGDIVSVSNICIVVDQT---SVSLHKSRLLSASDEK-----CF 570
Db 639 GVLGFWVNGSFIDSKSAD--GPRKCAQDFSHLYALNLRGNARTSGEERKQGGDGFDSG 696

571 SRPPVTFKFMNDSTIYKQQLGVNNEILLTTTYLETCEOENTYFYFOAKTDMYIYKNYEHKL 630
697 SRVTVAIIF-----FVKDKDAPNHTIP-----YYE-----VEDYLKREAKLNLANFENDL 742
631 TVPLSSITTLDTFIALNFTLLENVDFKVIELYTTDEKR--LSNVFDTIETMREYNYAQR 688
743 SVPFKEITPNDKGDWIN---QRNDDFEKLIPLKRDKKSKIFNAIFDLNS----- 788
689 VSLGRKDLLDLSTNRNOFVDAFGSLMDDLGAVGOTVNAVSGVATLFSISVTGFNFIKN 748
789 -NGVK-----TSRDPWVNFQ-----KTLMSQVQNCIDTVDNADLKRNERPRE 831
749 PF-----GGMLMIIVVIGVLFAYFLTKTKIYETAPIKMIYPEIDKLKEREGKSEIA 801
832 AFQRTAKDKGKKSADRYKHLNDREITTDKTIATW-----DGLANKLIKNEUL 881
802 PISEEELERIVIAWH 816
882 P--ESGMERVLALY 894

RESULT 38
US-09-732-180-7
; Sequence 7, Application US/09732180
; Patent No. US20020137702A1
; GENERAL INFORMATION:
; APPLICANT: Arkowitz, Robert A
; APPLICANT: Neim, Peter MA
; TITLE OF INVENTION: Yeast receptor
; FILE REFERENCE: DY0025.001AUS
; CURRENT APPLICATION NUMBER: US/09/732,180
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,699
; PRIOR FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 7
; LENGTH: 813
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-732-180-7

Query Match      2.5%; Score 115; DB 10; Length 813;
Best Local Similarity 19.2%; Pred. No. 1.2;
Matches 150; Conservative 107; Mismatches 267; Indels 256; Gaps 40;

QY 75 NIYGSPTFPYRVCASGVDFRQFDTHVCPDASDMVHSEGIILLYKQNIIPFMRVRYK 134
Db 148 NISVDPATHLWKLPQ---GAPFCVLINHLIPDSQIPVWSSDDLRIKKSVDVFLIAVKT 204

QY 135 Y-----RKVVTSTVYNGIYSDSITNQHFTFYKSTIEPHEKMDTIYQCFNSLRNTGGL 189
Db 205 QLNFDDENMFTISNVFSDNAQDLI-----KIIDVINKL-----L 238

QY 190 LTYVDR-----DDINNTVFLQPDV-----GVTPDKRYGSOPELYLEPGWFGSYR 235
Db 239 AEYSDASDSGGEDVNDV--QITDERSKVFREIIEETERYKVODLELMCK-----Y 288

QY 236 RRTTVNCELMD-----MFARSNPPDF-----FVTATGDT 265
Db 289 RQDLIEAENLSSEGIHLHLPNLNEIIDFORFLNGLECNINVPYRIGSVFIHAS--- 345

QY 266 VEMSPFWSEGDHDKNKHKEKFPVSVINNYKVVDYQNRGTVPGLKTRIFLD-----REY 321
Db 346 --LGPFNAYE-----PW---TIGQLTAIDLINKEAANKSSSLDPGFELOS 390

QY 322 LSWEKHLKNMSPCLTLWKAFYN---GIQTEHSGSVHFVANDITASFTTSKEDMKFNT 377
Db 391 L---KPIQRICKYPLLLKELIKTSPEYSKODPHGSSSSTSFNELLVAKTAMKELANQN- 446

QY 378 TYHCLNEEIKAEIKYAKVNSTHYSKYGDLYFKTDGGLYLVWQPLIQNRLLDAKNKLNN 437
Db 447 -----EAQRRAENIEHLEKLK---ERVGNWRGNLDAQGELLPHG---QVGVKDAENEKEY 496
```

```
Qy 438 ETY-----SRRSROAESTTDPMMEMTNGAGGEYSSSENSITVAQVQYAYD 483
Db 497 VAYLPEKIVFFFTTIDTKSDKQEKKS-----KFSTRKRSTSSNLSSSTT 542
Qy 484 NLRIRINILEDLSKAWCREQHRAALVWNELS-KINPTSVMSMIYNRPVSAKRI-GDVIS 541
Db 543 NLLSSINNRKQ-----NTLPLELKGVRVISEIYN--ISAPNTPGSTLI 584
Qy 542 VS-----NC-----IVVDQTSVSLHKSRLLSASDEKCFSRPP 574
Db 585 ISWSGRKESGFTLRYRSEEARNQWEKCLRLKTNEMNKQIHKLR-----DSOSSEN 637
Qy 575 VTFKFMNDSTYKQGLGVNNEILLTTT---YLETQCENTEYVFOAKTDMYIKNYEHLKT 631
Db 638 -----TDDSAIY-DYTGISTSPVNQSTQOQYDHRGSHSRHSSSTLSMMKN-NRVKS 690
Qy 632 VPLSSI-----TTLDTFI-----ALNFTLLENVDFKVE-----LYTRDEKRLSN 671
Db 691 GDLRSISSTSTLDSFNNLNGSPNTNPSLMSSDATKTPTFDVAIKLLYKSTE--LSE 748
Qy 672 VFDIETMFREYNYAQRV-----SGLRKDLDLSTNRNQFVDAFGSLM-----DDLGA 720
Db 749 PLIVNAQI-EYNDLLQKIISQIITSNLVAD--DVNISRLRYKDEGDFVNLNSDDDWGLV 805

RESULT 39
US-10-054-399A-29
; Sequence 29, Application US/10054399A
; Publication No. US20030175712A1
; GENERAL INFORMATION:
; APPLICANT: Arkowitz, Robert A
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES AND PROTEIN
; FILE REFERENCE: DYOUL3.1A2CPI
; CURRENT APPLICATION NUMBER: US/10/054.399A
; CURRENT FILING DATE: 2002-12-13
; PRIOR APPLICATION NUMBER: 09/169,474
; PRIOR FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: 08/951,141
; PRIOR FILING DATE: 1997-10-15
; PRIOR APPLICATION NUMBER: 09/529,106
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: US 09/732,180
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,699
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: PCT/GB98/03033
; PRIOR FILING DATE: 1998-10-08
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 813
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-054-399A-29

Query Match 2.5%; Score 115; DB 12; Length 813;
Best Local Similarity 19.2%; Pred. No. 1.2;
Matches 150; Conservative 107; Mismatches 267; Indels 256; Gaps 40;

Qy 75 NIYGSPSTFPYRVCASGVDFRFTQDHCVPDASDMVHSEGLLIYKQNIIPMFVRK 134
Db 148 NISVDPAATHLWKLFQO---GAPFCVLINHLPSQIPWSSDDLCKKSVYDFLIAVKT 204
Qy 135 Y-----RKVVTTSTVYNGIYSDSTINQHTFYKSIETPEWTEKMDTIYOCFNSRLNTG 189
Db 205 QLNFDDENMTFISNVFSDNAQDLI-----KLIIVINKL-----L 238
Qy 190 LTYVDR-----DDINTVFLQPD-----GVTPDVKRYGSPQELYLEPGWFGSYR 235
Db 230 AEYSDASGSGGDEVDNDV--QITDERSKVFREIETERYKVQDLELMCK-----Y 288
```

```
Qy 236 RRTTVNCELMD-----MEARSNPPPDF-----FVTATGDT 265
Db 289 RQOLIAEALSSSQIHLFPNLNEIIDFORFLNGLECNINPIRYQIRGCVFIHAS--- 345
Qy 266 VEMSPFWSGDEDDHENKHEKPFVSVINNYKVVDYQNRGTVPGLKTRIFLD-----REY 321
Db 346 --LGPENAYE-----PW---TIGQLTAIDLINKEAANKKSSSLLDPGFELOS 390
Qy 322 LSWEKHLKNMSYCPPLNLWAFYN-----GIQTEHSGSYHFVANDITASFTTSKEDMKE 377
Db 391 L---KPIORLCKYPLLLKELIKTSPESYKQDPHGSSSTSFNELLVAKTAMKELANQVN- 446
Qy 378 TYHCLNEELKAELEKKYKAVNTHSKYGDLYKFTDGLGLYLVWQPLIQLNLLDAKLN 437
Db 447 -----BAORRAENIEHLEK---ERVGNWRGNLDAQGELLPHG--QVCVKDAENE 496
Qy 438 ETY-----SRRSROAESTTDPMMEMTNGAGGEYSSSENSITVAQVQYAYD 483
Db 497 VAYLPEKIVFFFTTIDTKSDKQEKKS-----KFSTRKRSTSSNLSSSTT 542
Qy 484 NLRIRINILEDLSKAWCREQHRAALVWNELS-KINPTSVMSMIYNRPVSAKRI-GDVIS 541
Db 543 NLLSSINNRKQ-----NTLPLELKGVRVISEIYN--ISAPNTPGSTLI 584
Qy 542 VS-----NC-----IVVDQTSVSLHKSRLLSASDEKCFSRPP 574
Db 585 ISWSGRKESGFTLRYRSEEARNQWEKCLRLKTNEMNKQIHKLR-----DSOSSEN 637
Qy 575 VTFKFMNDSTYKQGLGVNNEILLTTT---YLETQCENTEYVFOAKTDMYIKNYEHLKT 631
Db 638 -----TDDSAIY-DYTGISTSPVNQSTQOQYDHRGSHSRHSSSTLSMMKN-NRVKS 690
Qy 632 VPLSSI-----TTLDTFI-----ALNFTLLENVDFKVE-----LYTRDEKRLSN 671
Db 691 GDLRSISSTSTLDSFNNLNGSPNTNPSLMSSDATKTPTFDVAIKLLYKSTE--LSE 748
Qy 672 VFDIETMFREYNYAQRV-----SGLRKDLDLSTNRNQFVDAFGSLM-----DDLGA 720
Db 749 PLIVNAQI-EYNDLLQKIISQIITSNLVAD--DVNISRLRYKDEGDFVNLNSDDDWGLV 805

RESULT 40
US-09-732-180-2
; Sequence 2, Application US/09732180
; Patent No. US20020137702A1
; GENERAL INFORMATION:
; APPLICANT: Arkowitz, Robert A
; APPLICANT: Nern, Peter MA
; TITLE OF INVENTION: Yeast receptor
; FILE REFERENCE: DYO25.001AUS
; CURRENT APPLICATION NUMBER: US/09/732,180
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,699
; PRIOR FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 844
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-732-180-2

Query Match 2.5%; Score 115; DB 10; Length 844;
Best Local Similarity 19.2%; Pred. No. 1.3;
Matches 150; Conservative 107; Mismatches 267; Indels 256; Gaps 40;

Qy 75 NIYGSPSTFPYRVCASGVDFRFTQDHCVPDASDMVHSEGLLIYKQNIIPMFVRK 134
Db 162 NISVDPAATHLWKLFQO---GAPFCVLINHLPSQIPWSSDDLCKKSVYDFLIAVKT 218
Qy 135 Y-----RKVVTTSTVYNGIYSDSTINQHTFYKSIETPEWTEKMDTIYOCFNSRLNTG 189
```

```

Db 219 QLNFDENFTISNFSDNAQDLI-----KIIVINKL-----L 252
QY 190 LTYVDR-----DDINMTVFLQPD-----GVTPDVKRYGSOPELYLEPGWFGSYR 235
Db 253 AEYSDASDSGGDEDVMDV--QITDERSKVFREIETERYKVQDLELMCK-----Y 302
QY 236 RRTTVNCELMD-----MFARSNPPDF-----FVTATGDT 265
Db 303 RQDLIEAENLSSEQIHLLFPNLNEIIDFORFLNGLECNINVPYRQIGSVFIHAS--- 359
QY 266 VEMSPWSGEDDHENKHEKDFVSVNNYKVDYQNRGTVPGLKTRIFLD-----REYVT 321
Db 360 --LGPFNAYE-----PW---TIGQLTAIDLINKEAANKKSSSLDGPFGLOSYI 404
QY 322 LSWKHLKNMNSCYPLTLWKAFYN---GIQTEHSGSYHFVANDITASFTTSKEDMKFPT 377
Db 405 L---KPIORLCKYPLLKELIKTSPEYSKQDPHGSSSTSFNELLVAKTAMKELANQVN- 460
QY 378 TYHCLNEEIKAEIEKKYAKVNSTHSGYKDYKFTDGGLYLVWQPLIQNRLLDAKNKLN 437
Db 461 ----EAORRAENIEHLEKLEK--ERVGNWRGFNLDAQGELLFHG--QVGKDAENEKEY 510
QY 438 ETY-----SRRSRQAESTTDPMMETGNGAGGEYSSENSITVAQVQYAYD 483
Db 511 VAYLFKIVFFTEIDDTKKSDQEKKS-----KFSTRKRSTSSNLSSTT 556
QY 484 NLRIRINILEDLSKAWCREQRAALVMNELS-KINPTSVMSMIYNRPVSAKRI-GDVIS 541
Db 557 NLLESINNSRKD-----NTLPLELKGRVYISEIYN--ISAPNTFGSLTI 598
QY 542 VS-----INVDQTSVLSLHKSRLLSASDEKCFSRPP 671
Db 705 GDLRSISSTSLDSSFNNGSPNTNPSLMSSTATKIPTFDVAIKLLYKSTE--LSE 762
QY 672 VFDIETMFREYNYAQRV-----SGLRKDLDLSTNRNQFVDAFGSLM-----DDLGA 720
Db 763 PLIVNAQI-EYNDLLQKIISQIITSNLVAD--DVNISRLRYKDDGDFVNLNSDDDWGLV 819

```

RESULT 41

```

US-10-054-399A-24
; Sequence 24, Application US/10054399A
; Publication No. US2003017512A1
; GENERAL INFORMATION:
; APPLICANT: Arkowitz, Robert A
; APPLICANT: Kern, Peter MA
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES AND PROTEIN
; TITLE OF INVENTION: SEQUENCES
; FILE REFERENCE: DYOUI3.1A2CP1
; CURRENT APPLICATION NUMBER: US/10/054,399A
; CURRENT FILING DATE: 2002-12-13
; PRIOR APPLICATION NUMBER: 09/168,474
; PRIOR FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: 08/951,141
; PRIOR FILING DATE: 1997-10-15
; PRIOR APPLICATION NUMBER: 09/529,106
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: US 09/732,180
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,699
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: PCT/GB98/03033
; PRIOR FILING DATE: 1998-10-08
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0

```

```

; SEQ ID NO 24
; LENGTH: 844
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-054-399A-24

```

```

Query Match      2.5%; Score 115; DB 12; Length 844;
Best Local Similarity 19.2%; Pred. No. 1.3;
Matches 150; Conservative 107; Mismatches 267; Indels 256; Gaps 40;

```

```

QY 75 NIYSPSTFPYRVCSAGVDFRQTDHVCDDASDMVHSEGIILYKONIIPMFRVRK 134
Db 162 NISVDPAHLWKLFQ---GAPFCVLINHILPDSQIPVSSDDLCKSVYDFLIAVKT 218
QY 135 Y-----RKVVTSTVYNGIYSDSITNQHTFYKSIETPETERKMDTIYOCFNSLRNLTCGNL 189
Db 219 QLNFDENFTISNFSDNAQDLI-----KIIVINKL-----L 252
QY 190 LTYVDR-----DDINMTVFLQPD-----GVTPDVKRYGSOPELYLEPGWFGSYR 235
Db 253 AEYSDASDSGGDEDVMDV--QITDERSKVFREIETERYKVQDLELMCK-----Y 302
QY 236 RRTTVNCELMD-----MFARSNPPDF-----FVTATGDT 265
Db 303 RQDLIEAENLSSEQIHLLFPNLNEIIDFORFLNGLECNINVPYRQIGSVFIHAS--- 359
QY 266 VEMSPWSGEDDHENKHEKDFVSVNNYKVDYQNRGTVPGLKTRIFLD-----REYVT 321
Db 360 --LGPFNAYE-----PW---TIGQLTAIDLINKEAANKKSSSLDGPFGLOSYI 404
QY 322 LSWKHLKNMNSCYPLTLWKAFYN---GIQTEHSGSYHFVANDITASFTTSKEDMKFPT 377
Db 405 L---KPIORLCKYPLLKELIKTSPEYSKQDPHGSSSTSFNELLVAKTAMKELANQVN- 460
QY 378 TYHCLNEEIKAEIEKKYAKVNSTHSGYKDYKFTDGGLYLVWQPLIQNRLLDAKNKLN 437
Db 461 ----EAORRAENIEHLEKLEK--ERVGNWRGFNLDAQGELLFHG--QVGKDAENEKEY 510
QY 438 ETY-----SRRSRQAESTTDPMMETGNGAGGEYSSENSITVAQVQYAYD 483
Db 511 VAYLFKIVFFTEIDDTKKSDQEKKS-----KFSTRKRSTSSNLSSTT 556
QY 484 NLRIRINILEDLSKAWCREQRAALVMNELS-KINPTSVMSMIYNRPVSAKRI-GDVIS 541
Db 557 NLLESINNSRKD-----NTLPLELKGRVYISEIYN--ISAPNTFGSLTI 598
QY 542 VS-----INVDQTSVLSLHKSRLLSASDEKCFSRPP 574
Db 599 ISWSGRKESGFTLRYRSEARNQWEKCLRDLTNEMNKQIHKKLR-----DSDSSFN--- 651
QY 575 VTFKPMNDSTIYKQGLVNNELLTTT---YLETQCENTEYVFOAKTDMYIKNYEHLKT 631
Db 652 ----TDDSAIY-DYTGISTSPVNSTQOQYDHRGSHSRHSSSTLSMMKN-NRVKS 704
QY 632 VPLSSI-----TTLDTFI-----ALNFTLLENVDFKVE-----LYTRDEKRLSN 671
Db 705 GDLRSISSTSLDSSFNNGSPNTNPSLMSSTATKIPTFDVAIKLLYKSTE--LSE 762
QY 672 VFDIETMFREYNYAQRV-----SGLRKDLDLSTNRNQFVDAFGSLM-----DDLGA 720
Db 763 PLIVNAQI-EYNDLLQKIISQIITSNLVAD--DVNISRLRYKDDGDFVNLNSDDDWGLV 819

```

RESULT 42

```

US-09-815-242-13693
; Sequence 13693, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.

```


Db 1465 ESVAESVASESVASVAVSDISDLTYTSEWSTSDSKIVASTVSPSSQSRSSIPMS 1524
Qy 395 AKVNSTHSGYGLKFKYKTDGGLYLWQPLIQNRLLDAKNLNE-----TYSRR----- 443
Db 1535 SDESSESR-----ESSGTL-----SENSDSIPTTFSTRYSPSGM 1563
Qy 444 -SRQAESTTDPMMETGNGAGGEYSSENSITVAQVAYDNLIRINNILEDLSKAWCR 502
Db 1564 SSRHYTNSSTSVSDWSSVAGDSESSVS-----ISESSESVTS 1606
Qy 503 EQHRAALVWNLKINPTSVMSIYNRPVSAKRIGDVISVNCIVVDQTSV---SLHKS 559
Db 1607 ESVAESVASE--SVTAVSDISDLTY-----TSEVSTSDSKIVPSTVSPSSQSRSS 1657
Qy 560 RLLSASDEKCFSPRPVTFKPMNDSTIYKGLGVNNEILLTTTLETCOENTEYFQAQTD 619
Db 1658 PIMSSDSESSRE-----SSGTL-----LSENSDSIPTTF-----STRYSPSGMS 1701
Qy 620 MYIYKN 625
Db 1702 SRHYTN 1707

RESULT 44
US-10-304-095-6
; Sequence 6, Application US/10304095
; Publication No. US20030134275A1
; GENERAL INFORMATION:
; APPLICANT: Long, David M.
; APPLICANT: Metz, Anneke M.
; APPLICANT: Love, Ruschelle A.
; FILE OF INVENTION: Telomerase Reverse Transcriptase (TERT) Genes
; CURRENT APPLICATION NUMBER: US/10/304,095
; CURRENT FILING DATE: 2002-11-26
; PRIOR APPLICATION NUMBER: US/09/417,485
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2184
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (330)...(335)
; OTHER INFORMATION: Xaa at position 330 = Leu or Ile;
; OTHER INFORMATION: Xaa at position 335 = Asp or Gly.
US-10-304-095-6

Query Match 2.5%; Score 114; DB 12; Length 2184;
Best Local Similarity 17.6%; Pred. No. 7.1; Mismatches 297; Indels 338; Gaps 46;
Matches 166; Conservative 140;
Qy 61 RGNENRDEEQNKNIYSPSTFYRVCSASGVDFRQTDHVCPCDASDMVHSEGLLI 120
Db 1031 RNNNNKKAQNEKNI-----DSENLEKKKKIYI 1061
Qy 121 YK-QNIIP---PMFRV-----RKVRKVVTSTV-----YNGI 148
Db 1062 YKIKILIEKRNPLKLSINHFISKKLRINWIPKKGLRPLNLTLNVPPIVQRIPEI 1121
Qy 149 YSDSITNQHTFYKSIPEPETERKMDT-----IYQCFNSRLN 184
Db 1122 LKSKSSSEFYFNILNUNLEREKDKNIKKRKYKNKFNPNVSLNINCFSLKCLGNMHHN 1181
Qy 185 TGG---NLLTYVDRDDINWTVLPQPDVGTVPVKRYGQPELYLBPQFWGSRRTVTN 241
Db 1182 NNSLFKNTLTGTELELKKWLH-----YLK-NWFKYKKRKKYIK 1222
Qy 242 CELMDMFARSNPPDFVATGDTVMGPFWSGEDDHENKMKHPW-----FVSVINNYK 296
Db 1223 NKL-----KNNKIYAYI-CIGD-----FSNCYEHLN--HNYLFKILKNFFDNINNFE 1267

Qy 297 VVDYQNRGTVPGLKTRIFLDREEVYTLSEKHLKN--MSYCPLTLTWKAFYNG----- 345
Db 1268 F-----IYLFKRSFRL-YNNKLNNSFLSYYPVVV-KSF--GLHYIRNLRE 1308
Qy 346 --IQTEHSGSVHFVANDITASFTTSKEDMKEFNTTYHCLNBEIEKAEIEK----- 392
Db 1309 LIIKSHLNDNHHFLNQMFK--TKSKSDLYFADSYKSL-----QVDRKDFWITIITVI 1360
Qy 393 KYAKVNSTHSGYGLKFKYKTDGGLYLWQPLIQNRLLDAKNLNNETYSRRSQAESTT 452
Db 1361 RYYVLNI-----YFSIKEFKLNRKNIFYQ-----IFQENQMGVYLSVRDKKRVENIK 1409
Qy 453 DPMEMTGNGAGGEYSSENSITVAQVAYDNLIRINN-----ILEDLKAWC 501
Db 1410 KWL-----NSMKKINHDEILESLSKNSSINNNKPMICTHNEQDTEKGNQTN 1458
Qy 502 REQHRAAL--VWNLKINPTSVMSIYNRPVSAKRIGDVISVNCIVVDQTSVSLHK-- 557
Db 1459 KEKHDIVIGPIYNNSPDSTTTTHSSNNYK-----GNNIHVSGDYKNDGL--LHKGN 1507
Qy 558 -SLRLLSASDEKCFSPRPVTFKPMNDSTIYKGLGVNNEILLTTTLETCOENTEYFQ 615
Db 1508 NSMNECYVKDICKNN-----NNNNNNNNNNNN-----SYNKLNCVTN-----N 1548
Qy 616 AKTDMYIKVNEHLKTVPLSSITLDTFIALNFTLLENVDFKVIETYTRDEKRLSNVFDI 675
Db 1549 SKNDIIKYHK-----TIDTDSKNHTYFKNKLNFEL-----DKIISNIYGL 1590
Qy 676 ETMFREYN-----YYA-----QVSGLRKDL-----LDLSTNRNQFVDA 709
Db 1591 PQGFSNLNLSLYAYLDKNEEPQNLSEKQINNKYFLANGTCYFNLSLILRIFDD 1650
Qy 710 FGLMDDLGVAGQTVNAVSGVATLPSISIVTGFINFNKPFGLMILIIIVIGULFAIYFL 769
Db 1651 F-----LFTLNKNIKIFKN-----LLKKKKITWGSNNIN 1679
Qy 770 TKTKIYETAPI---KWIYPEIDKLEREGKSEIAPISEEELERIVLAMHIHQO---N 821
Db 1680 SSKTKIFKPLIYKNDLLIYFNQYKQKKYK-----IKKKKIQSVNRKRIHQLNVAN 1734
Qy 822 SHMETKTRKDPKDSILTRAQNMRLKRSGY---SNLKNAESV 859
Db 1735 KKHSTSVQDKINKYINLIHPTIQKNDOSVLSNSIMNFERI 1775

RESULT 45
US-10-242-056-53
; Sequence 53, Application US/10242056
; Publication No. US20030113323A1
; GENERAL INFORMATION:
; APPLICANT: Ensign, Jerald C
; APPLICANT: Bowen, David J
; APPLICANT: Petell, James
; APPLICANT: Fatig, Raymond
; APPLICANT: Schoonover, Sue
; APPLICANT: French-Constant, Richard
; APPLICANT: Orr, Gregory L
; APPLICANT: Merlo, Donald J
; APPLICANT: Roberts, Jean L
; APPLICANT: Rocheleau, Thomas A
; TITLE OF INVENTION: Insecticidal Protein Toxins from
; TITLE OF INVENTION: Photorhabdus
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dowelanco
; STREET: 9330 Zionsville Road
; CITY: Indianapolis
; STATE: IN
; COUNTRY: US
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/242,056
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/063,615
FILING DATE: 18-MAY-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/395,497
FILING DATE: 28-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/007,255
FILING DATE: 06-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/608,423
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/705,484
FILING DATE: 28-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/743,699
FILING DATE: 06-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Borucki, Andrea T.
REGISTRATION NUMBER: 33651
REFERENCE/DOCKET NUMBER: 50301E
TELEPHONE: 317-337-4846
TELEPHONE: 317-337-4847
INFORMATION FOR SEQ ID NO: 53:
SEQUENCE CHARACTERISTICS:
LENGTH: 1844 amino acids
TYPE: amino acids
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-10-242-056-53

Query Match 2.5%; Score 113.5; DB 15; Length 1844;

Best Local Similarity 19.5%; Pred. No. 6;

Matches 119; Conservative 82; Mismatches 219; Indels 189; Gaps 34;

Qy	108	ASDMVHSEGLLIYKON----	IIPFMRVRKRYKVVTTSTVNGIY-----SDSITN-Q 156
Db	1118	STDAESLGLYCTGYQGEDTLVNFYMSQSSSYSDNNAPVTGLYIFADMSSDNMTNAQ 1177	
Qy	157	HTFYKSTIEPWETE--KMDTI-----	YOCFNSLRLAT---GG 187
Db	1178	ATNY-----WNSYPQEDTWADDSNKKVITRRVNNRYAEDYEIPSSVTSNSYSGWD 1232	
Qy	188	NLLTYVDRDDINMTVFLQPDVGTDPVKRYGSOPELYLEPGW----	PWGSYRRRTTVNCE 243
Db	1233	HSLTML-----	YGSVPNITFESAEDLRLSTNMAISIIHNGYAGTRRIQCN 1279
Qy	244	LMDMFARSNPPFFVATGDT-----	VEMSPFWSGE--DD-----HENKQHE-KPWFSV 290
Db	1280	LMQYASLGDKFIYDSFDDANRNLVPLFKGKDNSDDSDICINENPSSDDKKWYFS 1339	
Qy	291	VINNYKVVDYQNRGT--VPLGKTRIFLDREYTLSEKHLKNMSYCPLTLWKAFYNGIQT 348	
Db	1340	SKDKNKTADY-NGTQCIDAGTSN-----	KDFYNNLQ- 1370
Qy	349	EHGSGYHFVANDIPASFTTSKEDMKFNTTYHCLNEEIKAEIKYAKVNSTHSGYDGLK 408	
Db	1371	-----EIEVISVTGGYWSYKISNPIN-----INTGIDS-----	AKKVVTVKAGDDQ 1413
Qy	409	YEKTDGGLYLWQPL-----	IQNRLLDAKKK--LNNE-----TYSRRSRROAES 450
Db	1414	IFTADNSTYVPQPPAPSPFEEMIVQFNLLTIDCKNLNFDINQAHIEIDFTATAQDGRFLGA 1473	

Qy	451	TTDPMEMTCNGAGGE-----	YSSENSITVAQVOYAYDNLIRINNILEDLSKAWCREQH 505
Db	1474	ETP-IIPVTYKVLGTENVIALYSENGVQMIG-AY---	RRLNLTLFAQ----- 1518
Qy	506	RAALVWNELSKINP--	TSVMSMIYNRPVSAKRIGDVISVSNCTIVVDQTSVSLHKSRLLS 563
Db	1519	-----QLVSRANRGIDAVLSM-ETQNIQEPQLGAGTYVQ--	LVLDKYDESIH----- 1562
Qy	564	ASDEKCFSRPPVTFKFMNDS--	TIYKQOLGVNNEILL---TTTYLETQCENTEYFFQAK-- 617
Db	1563	-GTNKSFAIEYVDIFKENDSFVIYQOGLSETSQTVVVKVFLSYFIEATGNKNHLMVRAKYQ 1621	
Qy	618	---TDMYIY 623	
Db	1622	KETTDKILF 1630	

Search completed: October 28, 2003, 15:41:52

Job time : 94 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 15:29:40 ; Search time 392 Seconds
(without alignments)
2007.856 Million cell updates/sec

Title: US-10-055-364-24

Perfect score: 4547

Sequence: 1 MAGSLKRGSLVLMYLYQV.....KRGYSNLKNAESVEMLNTL 865

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 5728757 seqs, 909918778 residues

Total number of hits satisfying chosen parameters: 5728757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents_AA_Main:*

- 1: /cgn2_6/ptodata/2/paa/PCTUS_COMB.pep.*
- 2: /cgn2_6/ptodata/2/paa/US06_COMB.pep.*
- 3: /cgn2_6/ptodata/2/paa/US07_COMB.pep.*
- 4: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
- 5: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*
- 6: /cgn2_6/ptodata/2/paa/US081_COMB.pep.*
- 7: /cgn2_6/ptodata/2/paa/US082_COMB.pep.*
- 8: /cgn2_6/ptodata/2/paa/US083_COMB.pep.*
- 9: /cgn2_6/ptodata/2/paa/US084_COMB.pep.*
- 10: /cgn2_6/ptodata/2/paa/US085_COMB.pep.*
- 11: /cgn2_6/ptodata/2/paa/US086_COMB.pep.*
- 12: /cgn2_6/ptodata/2/paa/US087_COMB.pep.*
- 13: /cgn2_6/ptodata/2/paa/US088_COMB.pep.*
- 14: /cgn2_6/ptodata/2/paa/US089_COMB.pep.*
- 15: /cgn2_6/ptodata/2/paa/US090_COMB.pep.*
- 16: /cgn2_6/ptodata/2/paa/US091_COMB.pep.*
- 17: /cgn2_6/ptodata/2/paa/US092_COMB.pep.*
- 18: /cgn2_6/ptodata/2/paa/US093_COMB.pep.*
- 19: /cgn2_6/ptodata/2/paa/US094_COMB.pep.*
- 20: /cgn2_6/ptodata/2/paa/US095_COMB.pep.*
- 21: /cgn2_6/ptodata/2/paa/US096_COMB.pep.*
- 22: /cgn2_6/ptodata/2/paa/US097A_COMB.pep.*
- 23: /cgn2_6/ptodata/2/paa/US097B_COMB.pep.*
- 24: /cgn2_6/ptodata/2/paa/US098_COMB.pep.*
- 25: /cgn2_6/ptodata/2/paa/US099A_COMB.pep.*
- 26: /cgn2_6/ptodata/2/paa/US099B_COMB.pep.*
- 27: /cgn2_6/ptodata/2/paa/US100_COMB.pep.*
- 28: /cgn2_6/ptodata/2/paa/US101_COMB.pep.*
- 29: /cgn2_6/ptodata/2/paa/US102_COMB.pep.*
- 30: /cgn2_6/ptodata/2/paa/US103_COMB.pep.*
- 31: /cgn2_6/ptodata/2/paa/US104_COMB.pep.*
- 32: /cgn2_6/ptodata/2/paa/US106_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	4547	100.0	865	1	PCT-US00-18647-24

2	4547	100.0	865	26	US-10-055-364-24	Sequence 24, Appl
3	2473	54.4	793	26	US-10-055-364-55	Sequence 55, Appl
4	2438	53.6	830	26	US-10-055-364-45	Sequence 45, Appl
5	1953.5	43.0	874	12	US-08-804-439-16	Sequence 16, Appl
6	1953.5	43.0	874	17	US-09-301-390-16	Sequence 16, Appl
7	1953.5	43.0	874	17	US-09-338-326-16	Sequence 14, Appl
8	1949.5	42.9	808	12	US-08-804-439-14	Sequence 14, Appl
9	1949.5	42.9	808	17	US-09-301-390-14	Sequence 14, Appl
10	1949.5	42.9	808	17	US-09-338-326-14	Sequence 13, Appl
11	1940	42.9	824	26	US-10-055-364-43	Sequence 43, Appl
12	1937	42.6	845	12	US-08-804-439-94	Sequence 94, Appl
13	1937	42.6	845	17	US-09-301-390-94	Sequence 94, Appl
14	1937	42.6	845	17	US-09-338-326-94	Sequence 94, Appl
15	1927.5	42.4	829	14	US-09-053-601-3	Sequence 3, Appl
16	1927.5	42.4	829	23	US-09-831-000-13	Sequence 13, Appl
17	1918	42.2	823	26	US-10-055-364-37	Sequence 37, Appl
18	1916	42.1	792	26	US-10-055-364-42	Sequence 42, Appl
19	1902	41.8	808	26	US-10-055-364-38	Sequence 38, Appl
20	1898.5	41.8	791	26	US-10-055-364-41	Sequence 41, Appl
21	1871.5	41.2	818	26	US-10-055-364-44	Sequence 44, Appl
22	1848.5	40.7	857	7	US-03-360-107-113	Sequence 113, Appl
23	1848.5	40.7	857	12	US-08-804-439-18	Sequence 18, Appl
24	1848.5	40.7	857	17	US-09-301-390-18	Sequence 18, Appl
25	1844.5	40.6	829	17	US-09-338-326-18	Sequence 17, Appl
26	1839	40.4	856	8	US-10-055-364-46	Sequence 46, Appl
27	1839	40.4	856	8	US-08-471-913-103	Sequence 103, Appl
28	1839	40.4	856	8	US-08-475-668-103	Sequence 103, Appl
29	1839	40.4	856	8	US-08-484-223-103	Sequence 103, Appl
30	1839	40.4	856	8	US-08-484-223A-103	Sequence 103, Appl
31	1839	40.4	856	8	US-08-484-223D-103	Sequence 103, Appl
32	1839	40.4	856	8	US-08-485-551-103	Sequence 103, Appl
33	1839	40.4	856	8	US-08-487-266-103	Sequence 103, Appl
34	1839	40.4	856	8	US-08-487-266A-103	Sequence 103, Appl
35	1839	40.4	856	8	US-08-487-355-103	Sequence 103, Appl
36	1839	40.4	856	8	US-08-487-355A-103	Sequence 103, Appl
37	1839	40.4	856	13	US-08-919-600-103	Sequence 103, Appl
38	1839	40.4	856	19	US-09-502-445-103	Sequence 103, Appl
39	1839	40.4	856	28	US-10-267-682-103	Sequence 103, Appl
40	1839	40.4	856	28	US-10-267-748-103	Sequence 103, Appl
41	1792	39.4	849	12	US-08-804-439-17	Sequence 17, Appl
42	1792	39.4	849	17	US-09-301-390-17	Sequence 17, Appl
43	1792	39.4	849	17	US-09-338-326-17	Sequence 17, Appl
44	1787	39.3	874	12	US-08-804-439-15	Sequence 15, Appl
45	1787	39.3	874	17	US-09-301-390-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1
PCT-US00-18647-24
; Sequence 24, Application PC/TUS0018647
; GENERAL INFORMATION:
; APPLICANT: Patience, Clive
; TITLE OF INVENTION: Swine Gamma Herpesvirus DNA and Methods of Use
; FILE REFERENCE: 61750-299
; CURRENT APPLICATION NUMBER: PCT/US00/18647
; CURRENT FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: U.S. 60/142,736
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: U.S. 60/168,532
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 24
; LENGTH: 865
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Deduced amino
; OTHER INFORMATION: acid sequence of porcine gamma herpesvirus gpb
; OTHER INFORMATION: gene
PCT-US00-18647-24

```
Query Match      100.0%; Score 4547; DB 1; Length 865;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 865; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAGSLKRGSLALWLYQVALYSLAETGVTSPNTATWSTESPLTGHYTHDSSHGE 60
Db 1 MAGSLKRGSLALWLYQVALYSLAETGVTSPNTATWSTESPLTGHYTHDSSHGE 60

Qy 61 RGNENRDSBEONKNIYGSPTFPYRVCASGVGDVFRFQTDHVCDDASDMVHSEGIILLI 120
Db 61 RGNENRDSBEONKNIYGSPTFPYRVCASGVGDVFRFQTDHVCDDASDMVHSEGIILLI 120

Qy 121 YKQNIIPFMRVRYKRVKVTSTVYNGIYSDSIINQHTFYKSIETPWTETKMDTIYOCFNS 180
Db 121 YKQNIIPFMRVRYKRVKVTSTVYNGIYSDSIINQHTFYKSIETPWTETKMDTIYOCFNS 180

Qy 181 LRLNTGNNLLTYVDRDDINMTVFLQPDVGVTDPVKRYGSOPELYLEPGFWGYSYRRRTTV 240
Db 181 LRLNTGNNLLTYVDRDDINMTVFLQPDVGVTDPVKRYGSOPELYLEPGFWGYSYRRRTTV 240

Qy 241 NCELMDMFARSNPFPDFVTATGDTVEMSPFWSGEDDHENKMKHPWFSVINNYKVVDY 300
Db 241 NCELMDMFARSNPFPDFVTATGDTVEMSPFWSGEDDHENKMKHPWFSVINNYKVVDY 300

Qy 301 QNRGTVPGLKTRIFLDREEVYTLSEKHLKMSYCPCLTLWKAFYNGIOTHSYGHVFVAND 360
Db 301 QNRGTVPGLKTRIFLDREEVYTLSEKHLKMSYCPCLTLWKAFYNGIOTHSYGHVFVAND 360

Qy 361 ITASFTTSKEDMKFNTTYHCLNEEIKAIEKKYAKVNSTHSGYKGYDLKYFKTDGGLYLW 420
Db 361 ITASFTTSKEDMKFNTTYHCLNEEIKAIEKKYAKVNSTHSGYKGYDLKYFKTDGGLYLW 420

Qy 421 QPLIQNRLLDKAKNKLNNETYSRRSRROAESTTDPMMEMTNGAGGEYSSENSITVAQVQY 480
Db 421 QPLIQNRLLDKAKNKLNNETYSRRSRROAESTTDPMMEMTNGAGGEYSSENSITVAQVQY 480

Qy 481 AYDNLRIIRINNILEDLSKACREOHRALVWNLKSKINPTSVMSMIYNNRPVSAKRIGDVI 540
Db 481 AYDNLRIIRINNILEDLSKACREOHRALVWNLKSKINPTSVMSMIYNNRPVSAKRIGDVI 540

Qy 541 SVSNCIVVDQTSVSLHKSRLLSASDEKCFSRPPVTFKFMDNSTIYKQGLGVNNEILLTT 600
Db 541 SVSNCIVVDQTSVSLHKSRLLSASDEKCFSRPPVTFKFMDNSTIYKQGLGVNNEILLTT 600

Qy 601 TYLETQCENTEYYPQAKTDMYIKNYEHLKTVPLSSITTLDTFIALNFTLLENVDFKVI 660
Db 601 TYLETQCENTEYYPQAKTDMYIKNYEHLKTVPLSSITTLDTFIALNFTLLENVDFKVI 660

Qy 661 LYTRDEKRLSNVDFDIETMREYNYYAQRVSGRLKDLDDLSTNRNOFVDVAFGSLMDDLGA 720
Db 661 LYTRDEKRLSNVDFDIETMREYNYYAQRVSGRLKDLDDLSTNRNOFVDVAFGSLMDDLGA 720

Qy 721 GQTVNAVSGVATLFSISITGFINFIKNPFGGMLMIIVIGLVFAIYFLTKTKIYETAP 780
Db 721 GQTVNAVSGVATLFSISITGFINFIKNPFGGMLMIIVIGLVFAIYFLTKTKIYETAP 780

Qy 781 IKMIYPEIDKLKEREKSEAPISEEBLERIVLAMHIHQNSHMETKTRKDPKDSILTRA 840
Db 781 IKMIYPEIDKLKEREKSEAPISEEBLERIVLAMHIHQNSHMETKTRKDPKDSILTRA 840

Qy 841 QNMLRKESGYSLNKNAESVEMLNTL 865
Db 841 QNMLRKESGYSLNKNAESVEMLNTL 865
```

RESULT 2

```
US-10-055-364-24
; Sequence 24, Application US/10055364
; GENERAL INFORMATION:
; APPLICANT: Patience, Clive
; TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use
; FILE REFERENCE: #1750-379
```

```
; CURRENT APPLICATION NUMBER: US/10/055,364
; CURRENT FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: US/09/612,204
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US/60/142,736
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: US/60/168,532
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 24
; LENGTH: 865
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Deduced amino acid sequence of porcine gamma herpesvirus gpB gene
US-10-055-364-24
```

```
Query Match      100.0%; Score 4547; DB 26; Length 865;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 865; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAGSLKRGSLALWLYQVALYSLAETGVTSPNTATWSTESPLTGHYTHDSSHGE 60
Db 1 MAGSLKRGSLALWLYQVALYSLAETGVTSPNTATWSTESPLTGHYTHDSSHGE 60

Qy 61 RGNENRDSBEONKNIYGSPTFPYRVCASGVGDVFRFQTDHVCDDASDMVHSEGIILLI 120
Db 61 RGNENRDSBEONKNIYGSPTFPYRVCASGVGDVFRFQTDHVCDDASDMVHSEGIILLI 120

Qy 121 YKQNIIPFMRVRYKRVKVTSTVYNGIYSDSIINQHTFYKSIETPWTETKMDTIYOCFNS 180
Db 121 YKQNIIPFMRVRYKRVKVTSTVYNGIYSDSIINQHTFYKSIETPWTETKMDTIYOCFNS 180

Qy 181 LRLNTGNNLLTYVDRDDINMTVFLQPDVGVTDPVKRYGSOPELYLEPGFWGYSYRRRTTV 240
Db 181 LRLNTGNNLLTYVDRDDINMTVFLQPDVGVTDPVKRYGSOPELYLEPGFWGYSYRRRTTV 240

Qy 241 NCELMDMFARSNPFPDFVTATGDTVEMSPFWSGEDDHENKMKHPWFSVINNYKVVDY 300
Db 241 NCELMDMFARSNPFPDFVTATGDTVEMSPFWSGEDDHENKMKHPWFSVINNYKVVDY 300

Qy 301 QNRGTVPGLKTRIFLDREEVYTLSEKHLKMSYCPCLTLWKAFYNGIOTHSYGHVFVAND 360
Db 301 QNRGTVPGLKTRIFLDREEVYTLSEKHLKMSYCPCLTLWKAFYNGIOTHSYGHVFVAND 360

Qy 361 ITASFTTSKEDMKFNTTYHCLNEEIKAIEKKYAKVNSTHSGYKGYDLKYFKTDGGLYLW 420
Db 361 ITASFTTSKEDMKFNTTYHCLNEEIKAIEKKYAKVNSTHSGYKGYDLKYFKTDGGLYLW 420

Qy 421 QPLIQNRLLDKAKNKLNNETYSRRSRROAESTTDPMMEMTNGAGGEYSSENSITVAQVQY 480
Db 421 QPLIQNRLLDKAKNKLNNETYSRRSRROAESTTDPMMEMTNGAGGEYSSENSITVAQVQY 480

Qy 481 AYDNLRIIRINNILEDLSKACREOHRALVWNLKSKINPTSVMSMIYNNRPVSAKRIGDVI 540
Db 481 AYDNLRIIRINNILEDLSKACREOHRALVWNLKSKINPTSVMSMIYNNRPVSAKRIGDVI 540

Qy 541 SVSNCIVVDQTSVSLHKSRLLSASDEKCFSRPPVTFKFMDNSTIYKQGLGVNNEILLTT 600
Db 541 SVSNCIVVDQTSVSLHKSRLLSASDEKCFSRPPVTFKFMDNSTIYKQGLGVNNEILLTT 600

Qy 601 TYLETQCENTEYYPQAKTDMYIKNYEHLKTVPLSSITTLDTFIALNFTLLENVDFKVI 660
Db 601 TYLETQCENTEYYPQAKTDMYIKNYEHLKTVPLSSITTLDTFIALNFTLLENVDFKVI 660

Qy 661 LYTRDEKRLSNVDFDIETMREYNYYAQRVSGRLKDLDDLSTNRNOFVDVAFGSLMDDLGA 720
Db 661 LYTRDEKRLSNVDFDIETMREYNYYAQRVSGRLKDLDDLSTNRNOFVDVAFGSLMDDLGA 720

Qy 721 GQTVNAVSGVATLFSISITGFINFIKNPFGGMLMIIVIGLVFAIYFLTKTKIYETAP 780
Db 721 GQTVNAVSGVATLFSISITGFINFIKNPFGGMLMIIVIGLVFAIYFLTKTKIYETAP 780
```

```
QY 781 IKMIYPEIDKLEREGKSEIAPISEBELERIVLAMIHOQNSHMETKTRKDPKDSILTRA 840
|||||
Db 781 IKMIYPEIDKLEREGKSEIAPISEBELERIVLAMIHOQNSHMETKTRKDPKDSILTRA 840
|||||
QY 841 ONMLRKSGSYNLKNAESVEMLNTL 865
|||||
Db 841 ONMLRKSGSYNLKNAESVEMLNTL 865
|||||

RESULT 3
US-10-055-364-55
; Sequence 55, Application US/10055364
; GENERAL INFORMATION:
; APPLICANT: Patience, Clive
; TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use
; FILE REFERENCE: 61750-379
; CURRENT APPLICATION NUMBER: US/10/055,364
; CURRENT FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: US/09/612,204
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US/60/142,736
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: US/60/168,532
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 55
; LENGTH: 793
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Portion of Acelaphine herpesvirus.
US-10-055-364-55

Query Match 54.4%; Score 2473; DB 26; Length 793;
Best Local Similarity 57.4%; Pred. No. 1.2e-223;
Matches 468; Conservative 130; Mismatches 169; Indels 46; Gaps 8;

QY 74 KNIYGSPTFPYRVCSASGVDFRQTDHVCVPDASDMVHSEGIILLYKQNIIPMFVR 133
| | | | |
Db 1 KGHSDPSAPFRVCSASNIGDIFRFQTSKCPNTKDKHEHNEGILLIFKENVIPYKVR 60
| | | | |

QY 134 KYRKVVTTSTVNGIYSDITNOHTFYKSIETPEWETKMDTIYQCENSLRNTGGNLLTV 193
| | | | |
Db 61 KYRKIVTTSTIYNGIYADAVTQHVFSKSVPIYETRMDTIYQCYNSLDTVGGNLLVT 120
| | | | |

QY 194 DRDINNTPVLPQVDGTPDKRYGSOPELYLEPGFWGSRRTTIVNCELDMFARSNP 253
| | | | |
Db 121 DNDGSNMTVDLPQVDGLSNSVRRYHSQPEIHAEPGWLGGYRRRTTVNCEVTETDARVP 180
| | | | |

QY 254 PEDFFVTATGDTVEMSPFW-----SGEDDHENKHEKFWFVSVINNYKVDYQNRG 304
| | | | |
Db 181 PFRYFITNGDTIEMSPFWKAMNETETFGSDPRT-----LTVAKDYRVVDYKFRG 231
| | | | |

QY 305 TVPLGKTRIFLDREYTLGWKHLKMWSCPLTLWKAIFYNGIQTESHSGSYHFVANDITAS 364
| | | | |
Db 232 TPOQGHTRIFVDEEYTLWSAQOFRNISYCRWAHWKSFDAIKTEHKSRLHFVANDITAS 291
| | | | |

QY 365 FTTSKEDMKFNTTYTHCLNEEIKAEIKKYAKVNSTHSGYDLYKFKTDGGGLYVWQPLI 424
| | | | |
Db 292 FYTPNTQTRVLGKHVCLNNTIESELKSLAKVNDTHSPNGTAQYYLTNGLLWQPLV 351
| | | | |

QY 425 QNRLLDKAKKLN-----NETYSRRRQAEASTDPMWMTGNGAGGEYSSENSITV 475
| | | | |
Db 352 QOKLLDAGKLLDAVKKQNTTTTTTTRRRQRSVS-----SGIDDDVTAESTILL 403
| | | | |

QY 476 AQOQVADNLRIRINILEDLSKAWCREOHRALVNELSKINPTSVMSIYNRPVSAGR 535
| | | | |
Db 404 TQIQFAYDILRAQINNVLELSRAWCREOHRASLMMNELSKINPTSVMSIYGRPVSAKR 463
| | | | |

QY 536 IGDVISVNCIVVDQTSVLHKSRLLSA-SDEKCFSPRPVTFKFWNDSTIYKGOLGVNN 594
| | | | |
```

```
Db 464 IGDVISVSHCVVQDQSVSLHRSRMRVGRDKTHECYRPPVTFKPIINDSHLYKGOLGVNN 523
QY 595 EILLTTTLETQENTYFYFOAKTDMYIKYVNEHLKTVPLSSITTTLDTFIALNFTLENV 654
| | | | |
Db 524 EILLTTTAVEICHENTEHYFOGNNMYPYKNYHVKTMPVGVDVATLDTFFMVLNLTVENI 583
| | | | |
QY 655 DFKVIELYTRDEKRLSNVFDIETMFREYNYAQRVSGRLKDLDLSTNRNOFVDAFGSLM 714
| | | | |
Db 584 DFQVIELYSREKRMSTAFDIETMFREYNYTQRTVGLRRLDTLATNRNQFVDAFGSLM 643
| | | | |
QY 715 DDLGAVGQTVNAVSGVATLFSSTVTFPIFKNPPFGGMLMIIIVIGVLFAIYFLTKTK 774
| | | | |
Db 644 DDLGVGKTVLNAVSSVATLFSSTVSGIINPKNPPFGGMLLFGLIAAVITVILNRRAK 703
| | | | |
QY 775 IYETAPIKMIYPEIDKLEREGKSEIAPISEELERIVLAMIHOQNSHMETK----- 827
| | | | |
Db 704 RFAONPVOMI-PDIKTITSQREELQVDIPISKHELDRLMLAMHDYHASKQESKQDEEGS 762
| | | | |
QY 828 TRKDPKDSILTRAQNMLRKSGSYNLKNAESVE 860
| | | | |
Db 763 TTSGPAD--LNKAKNVLRRRAGYKPLKRTDSFE 793
| | | | |

RESULT 4
US-10-055-364-45
; Sequence 45, Application US/10055364
; GENERAL INFORMATION:
; APPLICANT: Patience, Clive
; TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use
; FILE REFERENCE: 61750-379
; CURRENT APPLICATION NUMBER: US/10/055,364
; CURRENT FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: US/09/612,204
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US/60/142,736
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: US/60/168,532
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 45
; LENGTH: 830
; TYPE: PRT
; ORGANISM: Alcelaphine herpesvirus
US-10-055-364-45

Query Match 53.6%; Score 2438; DB 26; Length 830;
Best Local Similarity 57.4%; Pred. No. 2.7e-220;
Matches 459; Conservative 129; Mismatches 166; Indels 46; Gaps 7;

QY 64 NENRDSBEON---KNIIYGSPTFPYRVCSASGVCDVFRFQTDHVCVPDASDMVHSEGIILLI 120
| | | | |
Db 42 SDNIIROQRNNTAKGIHSDPSAPFRVCSASNIGDIFRFQTSKCPNTKDKHEHNEGILLI 101
| | | | |

QY 121 YKQNIIPMFVRKYRKVKTSTVNGIYSDITNOHTFYKSIETPEWETKMDTIYQCENFNS 180
| | | | |
Db 102 PKENIVPVFKYRKIVKTSTIYNGIYADAVTQHVFSKSVPIYETRMMDTIYQCENFNS 161
| | | | |

QY 181 LRLNTGGNLLTYDRDDINMTVFLQPVGDVTPDKRYGSOPELYLEPGFWGSRRTTV 240
| | | | |
Db 162 LDVTVGNNLLVYTDNDGSNMTVDLPQVDGLSNSVRRYHSQPEIHAEPGWLGGYRRRTTV 221
| | | | |

QY 241 NCELMDMFARSNPFDFFVTATGDTVEMSPFW-----SGEDDHENKHEKFWFVS 291
| | | | |
Db 222 NCEVETDARAVPPFRYFITNIGDTIEMSPFWKAMNETETFGSDPRT-----LTV 272
| | | | |

QY 292 INNYKVDYQNRGTVPLGKTRIFLDREYTLGWKHLKMWSCPLTLWKAIFYNGIQTESH 351
| | | | |
Db 273 AKDYRVVDYKFRGTPOQGHTRIFVDEEYTLWSAQOFRNISYCRWAHWKSFDAIKTEH 332
| | | | |

QY 352 GSYHFVANDITASFTTSKEDMKFNTTYTHCLNEEIKAEIKKYAKVNSTHSGYDLYKFK 411
| | | | |
Db 333 KSLHFVANDITASFTTNTQTRVLGKHVCLNNTIESELKSLAKVNDTHSPNGTAQYYL 392
| | | | |
```

QY 412 TDGGLYLVWQPLQNRLLDAKKNLN-----NETYSRRSRRQAESTTDPMMMTGNG 462
Db 393 TNGGLLVWQPLQNRLLDAKGLLDVKKQNTTTTTTTRSRQRKRSVS-----SG 444
QY 463 AGGYSSENSITVAQVQVAYDNLRIRINNILEDLKAWCREQHRALVWNELSKINPTSV 522
Db 445 IDVYTAESTILLQIQFAYDTLRAQINNVLLELSRAWCREQHRASLWNELSKINPTSV 504
QY 523 MSMTYNRPVSAKRGDVISVNCIVDQTSVSLHKSRLLSA-SDEKCFRPPVTFKPMN 581
Db 505 MSSYGRPVSAKRICDVISHCVVVDQDSVSLHRSRMRVPGDKTHECYSRPPVTFKFIN 564
QY 582 DSTLYKGLGVNNEILLTTTILETCOENTYFYFOAKTDMYIKNYEHLKTVPLSSITL 641
Db 565 DSHLYKGLGVNNEILLTTTAVETICHENTHYFOGNNMYFYKYNRHKVTKMPGVGDVATL 624
QY 642 TFIALNFTLENDVFKVIELYTRDEKRLSNVDFIETMFREYNYAQRVSGRLKDLDLST 701
Db 625 TFWNLNLTVENIDFVIELYSREKRMSTAFDIETMFREYNYTORVTLGRDLTLAT 684
QY 702 NRNFVDAFGLMDDLGAQGVTVNAVSGVATLFPSSIVTGFINFIPKPFGLMLIIVIG 761
Db 685 NRNFVDAFGLMDDLGVGKTVLNAVSSVATLFPSSIVSGIINFIPKPFGLMLIIVIG 744
QY 762 VLFAIYELTKTIYETAPIKMIYPEIDKLKEREKSEIAPISEEELERIVLANHIQON 821
Db 745 VVITVILLNRKAKRFAQNPQVMIYDITKITSORELOVDPISKHELDRLMLAMHDYHAS 804
QY 822 SHMETK-----TRDKPKD 834
Db 805 KPESKQDEEGSTSGPAD 824

RESULT 5
US-08-804-439-16
; Sequence 16, Application US/08804439
; GENERAL INFORMATION:
; APPLICANT: Rose, Timothy M.
; APPLICANT: Bosch, Marnix L.
; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RPHV/KSHV
; TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,439
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Schiff, J. Michael
; REGISTRATION NUMBER: 40,253
; REFERENCE/DOCKET NUMBER: 29938-20002.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 874 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-804-439-16

Query Match 43.0%; Score 1953.5; DB 12; Length 874;
Best Local Similarity 46.1%; Pred. No. 2.2e-174;
Matches 403; Conservative 147; Mismatches 250; Indels 75; Gaps 18;

QY 11 VLALWLYQVALYSLSIAETGVTSPPNTATWSTESPLTGHYTHDSSHGGRGNNEURDSE 70
Db 10 VLCLWCV--AALLCQGAQEVVA-----ETTTT-----FATH-----RPEVAAE 46
QY 71 EQKNKIYGSPTSFPYRVCSASGV-GDVPFRQTDHVCPCDASDMVHSGILLIYKONIIIPM 129
Db 47 ENPANPF-----LPPFVCGASPTGGEIFRPPLEESCPNTEDKDHIEGIALIYKTNIVPV 101
QY 130 FVRKYRKVVVTTSTVYNGIYSDSITNQHTFYKSIIEPWETERKMDTIYQCNSLRNLG 189
Db 102 FNVRYKRIKMTSTTIYKGSSEDAITNQHTRSYAVPLYEQVMDHYIYQCFSAVQVNEGGHV 161
QY 190 LTYVDRDDINMTVPLOPVGVTPDVKRYGSOPELYLEPCFWGSGYRRRTTVNCELMDMFA 249
Db 162 NTYVDRDGNWETAFLKPADGLTSSITRYOSQPEVYATPRNLLMSYTRTRTVNCEVTEMSA 221
QY 250 RSNPPRDFEFTATGDTVEMSPFWSGEDDHENKMKHKEPWFVSVINNYKVVDYQNRGTVP 309
Db 222 RSMKPEFFVTSVGDITFEMSPFLKENGTEPEKILKRPHSIQLLKNYAVTKY---GVGLG 277
QY 310 K-----TRIFDLREBYTLSEKHLKNMSYCYPLTLWKAFYNGIQTEHSGSYHFVANDITAS 365
Db 278 QADNATRFPAIFGDIYSLSKWATTENSICYDLILWKGFSAIQTQHNSSLHFIANDITAS 337
QY 366 TTSKEDMKFNTTYHCLNEEIKAEIKYAKVNSHYSKYGDLYKFTDGLGLYLWVQPLIQ 425
Db 338 STPLEBEANFNETFKCIMWNTQOEIIOKLKUEKTHRPNGTAKVYKTTGNLXIVWQPLIQ 397
QY 426 NLLDAKNKLN-----ETYSRRSRQAESTTDPMMMTGNGAGEYS--ENS 472
Db 398 IDLLDTHAKLYNLTNATASPTSTPTSPRRRDRTSS-----VSGGNGNDNSTKEES 450
QY 473 ITVAQVQAYDNLRIIRINNILEDLSKAWCREQHRALVWNELSKINPTSVMSIYNRPVS 532
Db 451 VAASQVQAYDNLRLKRSINRVLGELSRAWCREQYRASLWVYELSKINPTSVMSAIYGRPVS 510
QY 533 AKRIGDVISVNCIVDQTSVSLHLSRLLSASDEKCFRPPVTFKFMNDSTIYKQGLV 592
Db 511 AKLIGDVSVDICISVDQKSVFVHKNMK-VPGKEDLCYTRPVVGVKFXINGSELFAGQLGP 569
QY 593 NNEILLTTLTLETQENTYFYFOAKTDMYIKNYEHLKTVPLSSIITLDTFIALNFTLLE 652
Db 570 RNEIVLSTSQVEVCQHSCEHYFOAGNQMYKYDYVYVSTNLNLTIDITLMTITLNLSLVE 629
QY 653 NVDFKVIELYTRDEKRLSNVDFIETMFREYNYAQRVSGRLKDLDD-LSTNRNQVDAFG 711
Db 630 NIDFKVIELYSKTEKLSNVDFIETMFREYNYVYQNLNGLRKDLDDSDIDGRDSPFIOTLG 689
QY 712 SLMDLGAQGVTVNAVSGVATLFPSSIVTGFNFINKPFGGLMLIIVI-VULFAIYFLT 770
Db 690 DIMQDLGTIGKVVVNAVSGVSLFGSIVSGVISFPPKPNPFGGLMLIIVLIITAGVVVVLFW 749
QY 771 KTKIYETAPIKMIYPEIDKLKEREKSEIAPISEEELERIVLANHIHOONSHMETKTRK 830
Db 750 RRSIY-SAPIRMLYPGVERAAQEPGAH---PVSEDIIRNIIIMGHQFOQORABEEARR 805
QY 831 DPK-----DSILTRAQNMLRKR---SGYSLNK 854
Db 806 EEEVKGKRLTFEVRDTSATSVLRRRRGGGGYQRLQ 840

RESULT 6
US-09-301-390-16
; Sequence 16, Application US/09301390
; GENERAL INFORMATION:

APPLICANT: Rose, Timothy M.
APPLICANT: Bosch, Marnix L.
APPLICANT: Strand, Kurt
TITLE OF INVENTION: GLYCOPROTEIN B OF THE RFHV/KSHV
SUBFAMILY OF HERPES VIRUSES
NUMBER OF SEQUENCES: 100
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morrison & Foerster
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1018

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/301,390
FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/720,229
FILING DATE: 26-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Schiff, J. Michael
REGISTRATION NUMBER: 40,253
REFERENCE/DOCKET NUMBER: 29938-20002.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 813-5600
TELEFAX: (415) 494-0792
TELEX: 706141

INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 874 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-301-390-16

Query Match 43.0%; Score 1953.5; DB 17; Length 874;
Best Local Similarity 46.1%; Pred. No. 2.2e-174;
Matches 403; Conservative 147; Mismatches 250; Indels 75; Gaps 18;

QY 11 VLALWYLQVALYSLSIAETGVTSPNTATWSTESPLTGHGYTHDSHGGRGNENRDS 70
DB 10 VLCLMCV--AALLCOGVAQVVA-----ETTP----FATH-----RPEVVAE 46
QY 71 EQKNKIYGPSFPYRVCSASGV-GDVERFQTDHVCPSDASDMVHSEGIILYKQNIIPFM 129
DB 47 ENPAMPF-----LPRVCGASTGGEIRFPPLSESCPNTEKDHIEGIALIYKTNIVPV 101
QY 130 FVRVRYKRVKVTSTVYNGIYSDISINQHTFYKSIPEWETEKMDTIYQCFNSLRNLGTGNL 189
DB 102 FNVRYKRMSTTIYKGNSEDAINQHTRSVAVPLYEVQMDHYQCFSAVQVNEGHRV 161
QY 190 LTYVDRDINMTVFIQPDVGTVPVKRYGSGQPELYLEPGFWGFSYRRRTTNCCLMDMFA 249
DB 162 NTYYDRDGNWETAFLKPADGLTSSITRYQSQPEVATPRLILWSYTRRTTNCVEVTEMSA 221
QY 250 RSNPPEFVATGTDTVMSPWSEDDHKNHKEKHPFVSVINNYKVVDYQNRGTVPGLG 309
DB 222 RSMKPEFVTSVGTIEMSPKENGTEPEKILKRPHSIQLLKNYAVTKY-----GVGLG 277
QY 310 K-----TRFLDREBYTLGWKHLKMWSCPLTLWKAIFYNGIQTEHSGSYHVFVANDITASF 365
DB 278 QADNATRFATPGDYSLSKWKATTENSSYCDLLWKGFSNAQTQNSLSLHFIANDITASF 337
QY 366 TTSKEDMKEFNITYHCLNEEIKAEIERKYAKVNSTHSGYGLKYFKPTDGGLYVMQPLIQ 425
DB 338 STPLEEAFNETFKCIWNNTQEEIQKLEKEKTHRPNGTAKVYKTTGNLYIVWQPLIQ 397

QY 426 NRLLDANKNLNN-----ETYSRRSRROAESTTDDMMEMTNGAGGEYSS-ENS 472
DB 398 IDLLDTHAKLYNLNATASPTSTPTTSRRRRRTSS-----VSGGNGNDNSTKEES 450
QY 473 ITVAQVQVAYDNLRIIRINNILEDLSKAWCREQRAALVNLKSLINPTSVMSIMYNRPV 532
DB 451 VAASQVQFAYDNLRSINRVLGELSRACREYRASLMWYELSKINPTSVMSAIYGRPV 510
QY 533 AKRIGDVISVNSCIWVDTOSVLSHLSRLLSASDEKCFSPPTFKWNDSTIYKGOLGV 592
DB 511 AKLIGDVVVSDCISVDOKSVFVHKNMK-VPGKEDLCYTRPVVGFKFNGSELFAGQLGP 569
QY 593 NNEILLTTLTETQENTYVFOAKTDWYIKNYEHLKTVPLSSITLDTFIALNFTLLE 652
DB 570 RNEIVLSTSQVEVCHSEHYFOAGNQMYKYVYVYSTLNLTDIPLHTMITLNLISLVE 629
QY 653 NVDPKVIELYTRDEKRLSNVFDIETMPREYNVYVYQAVSGRLRKOLLDD-LSTNRNQFVDAFG 711
DB 630 NIDPKVIELYSKTEKRLSNVFDIETMPREYNVYVYQNLNGLRKOLLDDSIDHGRDSFIOTLG 689
QY 712 SLMDDLGAVGOTVNVAVSVATLFSISIVTGFINIKNPFQGMMLIIVVI-GVLFAYFLT 770
DB 690 DIMQDLGTIGKVYVNVASGVFSLSFGSIVSGVISFKNPFGGMLLIIVIIAGVVVYVLFMT 749
QY 771 KKTIVETAPIKMIYPEIDKIKEREKSEIAPISSEELERIVLAMHIHQNSHMETKTRK 830
DB 750 RRSRIY-SAPIRMLYPGVERAAQEPGAH---PVSEDIQIRNLMGMHGFQQRORAEER 805
QY 831 DPK-----DSILTRAQNMRLKX---SGYSNLK 854
DB 806 EEEVKGKRTLFEVIRDSATSVLRRRGGGYQRLQ 840
RESULT 7
US-09-338-326-16
; Sequence 16, Application US/09338326
; GENERAL INFORMATION:
; APPLICANT: Rose, Timothy M.
; APPLICANT: Bosch, Marnix L.
; APPLICANT: Strand, Kurt
; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RFHV/KSHV
; SUBFAMILY OF HERPES VIRUSES
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/338,326
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/720,229
; FILING DATE: 26-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Schiff, J. Michael
; REGISTRATION NUMBER: 40,253
; REFERENCE/DOCKET NUMBER: 29938-20002.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:

```
;
; LENGTH: 874 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-338-326-16

Query Match
  43.0%; Score 1953.5; DB 17; Length 874;
Best Local Similarity 46.1%; Pred. No. 2.2e-174; Indels 75; Gaps 18;
Matches 403; Conservative 147; Mismatches 250;

Qy 11 VLALWLYQVALYSLSAETGVTSPPTATWSTESPLTGHYGDTHSSHGRGNENRDS 70
Db 10 VLCLWCV--AALLCQVAQEVVA-----ETTP-----PATH-----RPEVVAE 46

Qy 71 EQNKNIYGSSTFYRVCSAGV-GDVFRTQDTHVCPDASDMVHSEGILLIYKQNIIPFM 129
Db 47 ENPANPE-----LPRVCGASPTGGEIFRPPLSESCPTEDKOHIEGIALIYKTNIPVY 101

Qy 130 FRVKYKRVKVTSTVYNGIYSDITNOHTFYKSTIEPTEKMDTIYOCFNSLRINTGSL 189
Db 102 FNVKRYKRMKTTIYKGSBEDAITNOHTRSYAVPLVEYQMDHYOCFSAVQNEGGHV 161

Qy 190 LTYVDRDDINMTFLQPDVGTDPVKRYGSOPELYLEPGFWGYSYRRRTTVNCELMDMFA 249
Db 162 NTYVDRDGNWETAFLKPADGLTSSITRYQSOPEYATPRNLLNSYTRTRTVNCEVTEMSA 221

Qy 250 RSNPPDFVVTATGDTVEMGSPFWGEGEDDHENKMKHKEKFWFSVINNYKVQYQNGTVPLG 309
Db 222 RSMKPFDFVTSVGDITIEMSPFLKENGTEPEKILKRPHS1QLLNKNAVTKY---GVGLG 277

Qy 310 K----TRIFLDREBYTILSWEKHLNMSYCPILTAKAFYNGIOTHSYHFVANDITASF 365
Db 278 QADNATFAIFGYSLSWKATTENSSYCDLILWKGFSAIQTQHSLSLHFANDITASF 337

Qy 366 TTSKEDMKFNTTYHCLNEEKIKAEIKYAKVNSYKSGDKLYFKTDGGLYLWVQPLIQ 425
Db 338 STPLEEAEANFETFKIWNNTQBEIQKKEVEKTHRPNGTKAVKYKTTGNLYIVWQPLIQ 397

Qy 426 NRLLDKKNLNN-----ETYSRRSRROAESTTDPMMEMTNGAGGEYSS-ENS 472
Db 398 IDLLDTHAKLYNLNATASPTSTPTSPRRRRRTSS-----VSGGNGNGNSTKEES 450

Qy 473 ITVAQVQAYDNLRIIRNNILLEDLSKAWCEQHOAAALVWNLKINPTSVMSMTYNRPV 532
Db 451 VAASQVQFAYDNLKRSINRVLGELUSRAWCEQYRASLWYELSKINPTSVMSALYGRPV 510

Qy 533 AKRGDVISVNCITVVDQTSVSLHSLRLLSASDEKCFSRPPVTFKPMNDSTIYKQGLGV 592
Db 511 AKLIGDVVSVDCLISVDQSKSVFVHKMK-VPGKEDLCYTRPVVGFKEINGSELPAGQLGP 569

Qy 593 NNEILLTTLTLETQENETYYFOAKTDMYIKVNYEHLKTVPLSSITLDTFIALNFTLE 652
Db 570 RNEIVLSTSQVEVCQHSCEHYFOAGNQMYKYDYVYVSTLNLTDIPTLHTMITLINSLVE 629

Qy 653 NVDEKVELITRDEKRLSNVPDIETMREYNVYQAVSGLRKLDD-LSTNRQFVDAPG 711
Db 630 NIDPKVIELYSKTEKRLSNVPDIETMREYNVYQNLNGLRKLDLSDSIDHGRDSFIQTLG 699

Qy 712 SLMDLDGAVGQTVNAVSGVATLFSSTVTFINFKNPPFGMLMIIVVI-GVLFAYIFLT 770
Db 690 DIMQDLGTIGKVVNVASGVFSLFGSIVSGVIFSKPFGGMLLIVLIIAGVVVVYLFMT 749

Qy 771 KTKIYETAPIKMTYPIEDKLKEREGKEIAPISEEELERIVLAWHTHQONSHMETKTR 830
Db 750 RRSRIY-SAPIRMLYPGVGAQAQPPGAH---PVSEDDQIRNLIMGMHQFQQRAREEBARR 805

Qy 831 DPK-----DSLTRAQNMRLKX---SGYSNLK 854
Db 806 EEVKGKRTLFEVIRDSATSVLRRRRGGGGYQLQ 840

RESULT 8
```

```
US-08-804-439-14
; Sequence 14, Application US/08804439
; GENERAL INFORMATION:
; APPLICANT: Rose, Timothy M.
; APPLICANT: Bosch, Marnix L.
; APPLICANT: Strand, Kurt
; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RPHV/KSHV
; TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; APPLICATION DATA:
; CURRENT APPLICATION NUMBER: US/08/804,439
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Schiff, J. Michael
; REGISTRATION NUMBER: 40,253
; REFERENCE/DOCKET NUMBER: 29938-20002.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 808 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-804-439-14
```

```
Query Match 42.9%; Score 1949.5; DB 12; Length 808;
Best Local Similarity 46.6%; Pred. No. 4.6e-174;
Matches 373; Conservative 159; Mismatches 236; Indels 33; Gaps 10;

Qy 71 EQNKNIYGSSTFYRVCSAGVGDVFRTQDTHVCPDASDMVHSEGILLIYKQNIIPFM 130
Db 28 EKNTQAIQYQYKRVCSASTTGLFRDLDRTCPSTEDKVHKEGILLIYKKNIVPIYF 87

Qy 131 RVKRYKRVKVTSTVYNGIYSD--SITNOHTFYKSIPEWTEKMDTIYOCFNSLRINTGN 188
Db 88 KVRRYKKTITSVRIFNGWTRREGVAITNKLWELGRAVPKYEIDIMDKTYQCHNCMQIEVNGM 147

Qy 189 LTYVDRDDINMTFLQPDVGTDPVKRYGSOPELYLEPGFWGYSYRRRTTVNCELMDMF 248
Db 148 LNSYDRDGNKNTVDUKPDGLTGALTAYISQKVPADPGWLMTGTYTRTTVNCIELVDMF 207

Qy 249 ARSNPDPDFVTATGDTVEMSPFWGEGEDDHENKMKHKEKFWFSVINNYKVQYQNGTVPL 308
Db 208 ARSADPTTYFTALGDTVEVSPFCVDVNSCPNATDVL--VQIDLNHTVVVDYGNRATSOQ 265

Qy 309 GKTRIFLDREBYTILSWEKHLNMSYCPILTAKAFYNGIOTHSYHFVANDITASFTTS 368
Db 266 HKKRIFAHTLDYSVSEAVNKSASVCMVFWKSFQRAIQTEHDLTYHFIAEITAGFSTV 325

Qy 369 KEDMKFENTTYHCLNEEKAEIEKVKYAKVNSYKSGDKLYFKTDGGLYLWVQPLIQNRL 428
Db 326 KEPLANTSDYNCLMTHINTTLEDKJARVNNHTPNGTAEYQTEGMLVWVQPLIAEL 385

Qy 429 LDKNKLNN-----TYSRRSRROAESTTDPMMEMTNGAGGEYSSSENSITVAQVQYA 481
Db 386 EEAMLEATTSPVTPSAPTSSRSKRAIRSIRDV-----SAG-----SENNVFLSQIOYA 434
```

```
QY 482 YDNLIRINNIILEDLSKAWCREQRAALVWNLSEKINPTSVMSIYNRPVSAKRIGDVIS 541
Db 435 YDKLRQSINNVLBELAITWCREQVQRTWVYIAKINPTSVMTAIYKPSRKALGDVIS 494
QY 542 VNCIVVDQTSVSLHKSRLLSASDEKCFRPPVTFKFMNDSTIYKQGLGVNNEILLTTT 601
Db 495 VTECINVDQSSVSIHSLK--TENNDICYSRPPVTFKFMNSQLFKQGLGARNEILLSES 552
QY 602 YLETCQENTYVFOAKTDMYIKNYEHLKTVPLSSITLDTFIALNFTLLENVDFKVEL 661
Db 553 LVENCHQNAETFTAKNETYHFKNYVHVELPVNNISTLDTFLALNFTFIENIDFKAVEL 612
QY 662 YTRDEKRLSNVFDIETMFREYNYAQRVSGRLKDLIDS--TNRNQFVDAFGSLMDDLGA 719
Db 613 YSSGERKLANVFDLETMFREYNYAQSISGLRKD-FDNSQRNRRDRIIQDFSEILADLGS 671
QY 720 VQQTVMNAVSGVATLPSISIVTGFINFTKNPFGGMLMIIVIGVLFAIYFLTKTKIYETA 779
Db 672 IGVIVNVASGAFSLFGGIVTGILNFKNPLGGMFTLLIGAVIILVILLVRRTNMMSQA 731
QY 780 PIKMIYPEIDKLKEREGKSEIAPISEEELERIVLAWHIHQONSHMETKTRKDPKDSILTR 839
Db 732 PIRMIYPDVEK-----SKSTVTPMEPETIKOILLGHNMQOAYKKKEQRAARPSIFRQ 786
QY 840 -AQNMLRKRSYGNLKNAESV 859
Db 787 AAEFLRKESGYKQISTEDKI 807

RESULT 9
US-09-301-390-14
; Sequence 14, Application US/09301390
; GENERAL INFORMATION:
; APPLICANT: Rose, Timothy M.
; APPLICANT: Bosch, Marnix L.
; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RPHV/KSHV
; TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Morrison & Foerster
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/301.390
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/720,229
; FILING DATE: 26-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Schieff, J. Michael
; REGISTRATION NUMBER: 40,253
; REFERENCE/DOCKET NUMBER: 29938-20002.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 808 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
```

```
MOLECULE TYPE: protein
US-09-301-390-14
Query Match 42.9%; Score 1949.5; DB 17; Length 808;
Best Local Similarity 46.6%; Pred. No. 4.6e-174;
Matches 373; Conservative 159; Mismatches 236; Indels 33; Gaps 10;
QY 71 BONKNIYGSPTFYRVCSASGVDFRFQDTHVCYPDASDMVHSEGLIILYKONILPMPF 130
Db 28 EKNTQAIYQEFYKRVCSASTTGELFRFDLDRTCPSTEDKVHREGILLVYKKNIYVIPF 87
QY 131 RVRKYRKVVTSTVYNGIYSD--SITNQHTFYKSIEMPETEKMDTIYQCFNSLRANTGN 188
Db 88 KVRKYKITTIVRIFNGWTRGVAITNKWELSRAPVKEIDIMDKTYQCHNCMQIENVGM 147
QY 189 LLTYVDDDDINMTVFLQPDVCTPDVKRYGSQPELYLEPGHFWGFSYRRRTTNCCELMDF 248
Db 148 LNSYVDRDGNKNTVDLRKPDGLTGAI TRYISQPKFADPDGMLWGTYTRTTVNCIEIDMF 207
QY 249 ARSNPPDFFTATGDTVMSPFWSGEDDHENKHEKPFVSVINNVKVDYQNRGTVPL 308
Db 208 ARSADPTFYFTALGDIIVEVSPFCDDVNSCPNATDVL--VQIDLNTVVDYGNRATSOQ 265
QY 309 GKTRIFLDREBYTSLWEKHLKNMSYCPULTLWKAIFYNGIQTEHSGSYHFVANDITASF 368
Db 266 HKRIFAHTLDYSVSEAVNKSASVCSMVEFKSFQRAIQTEHDLTYHFIAEITAGFSTV 325
QY 369 KEDMKENPTYHCLNEIEKAEIEKKAIVNSTHSGYGLKYFKTDGGLYLVWQPLIQNRL 428
Db 326 KEPLANFTSYNCLMTHINTTLEDKIARVNNTHTPNGTAEYIQTGEGMILVWQPLIAEL 385
QY 429 LDKAKNLNE-----TYSRRSRQAEISTTDPMMENTGNGAGGEYSSENSITVAQOYA 481
Db 386 BEAMLEATTSPVTPSAPTSSRSKRAIRSDV-----SAG-----SENNVFLSQIOYA 434
QY 482 YDNLIRINNILEDLSKAWCREQRAALVWNLSEKINPTSVMSIYNRPVSAKRIGDVIS 541
Db 435 YDKLRQSINNVLBELAITWCREQVQRTWVYIAKINPTSVMTAIYKPSRKALGDVIS 494
QY 542 VNCIVVDQTSVSLHKSRLLSASDEKCFRPPVTFKFMNDSTIYKQGLGVNNEILLTTT 601
Db 495 VTECINVDQSSVSIHSLK--TENNDICYSRPPVTFKFMNSQLFKQGLGARNEILLSES 552
QY 602 YLETCQENTYVFOAKTDMYIKNYEHLKTVPLSSITLDTFIALNFTLLENVDFKVEL 661
Db 553 LVENCHQNAETFTAKNETYHFKNYVHVELPVNNISTLDTFLALNFTFIENIDFKAVEL 612
QY 662 YTRDEKRLSNVFDIETMFREYNYAQRVSGRLKDLIDS--TNRNQFVDAFGSLMDDLGA 719
Db 613 YSSGERKLANVFDLETMFREYNYAQSISGLRKD-FDNSQRNRRDRIIQDFSEILADLGS 671
QY 720 VQQTVMNAVSGVATLPSISIVTGFINFTKNPFGGMLMIIVIGVLFAIYFLTKTKIYETA 779
Db 672 IGVIVNVASGAFSLFGGIVTGILNFKNPLGGMFTLLIGAVIILVILLVRRTNMMSQA 731
QY 780 PIKMIYPEIDKLKEREGKSEIAPISEEELERIVLAWHIHQONSHMETKTRKOPKDSILTR 839
Db 732 PIRMIYPDVEK-----SKSTVTPMEPETIKOILLGHNMQOAYKKKEQRAARPSIFRQ 786
QY 840 -AQNMLRKRSYGNLKNAESV 859
Db 787 AAEFLRKESGYKQISTEDKI 807

RESULT 10
US-09-338-326-14
; Sequence 14, Application US/09338326
; GENERAL INFORMATION:
; APPLICANT: Rose, Timothy M.
; APPLICANT: Bosch, Marnix L.
; APPLICANT: Strand, Kurt
; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RPHV/KSHV
; TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
```

```
;
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/338,326
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/720,229
; FILING DATE: 26-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Schiff, J. Michael
; REGISTRATION NUMBER: 40,253
; REFERENCE/DOCKET NUMBER: 29938-20002.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 808 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-338-326-14

Query Match 42.9%; Score 1949.5; DB 17; Length 808;
Best Local Similarity 46.6%; Pred. No. 4.6e-174; Indels 33; Gaps 10;
Matches 373; Conservative 159; Mismatches 236;

Qy 71 EQNKNIYGPSPTFFRYVCSASGVDFRQDTHVCPDASDMVHSEGIILYKQNIIPFMF 130
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 28 EKNKTAIYQYFYKRVCSASTTGELFRFDLDCPTSTEDKVHKEGILLVYKKNIVYIF 87

Qy 131 RVRYKRVVTTSTVYNGIYSD--SITQHTPYKIEPWETKMDTIYOCFNSRLNTGNN 188
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 88 KVRYYKKTTSVRIFNGWTREGVAITNKLWSRAVPKYEIDIMDKTYQCHNCMOIEVNGM 147

Qy 189 LLTYVDRDDINMTVFLQPDGVTPDVKRYGSOPELYLEPGWFGSYRRRTTVCNCLDMF 248
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 148 LNSYDRDGNKNTVDLPVQDGLTGATRYISQPKVFADPGMLWGTGYTRITVNCIEVDMP 207

Qy 249 ARSNPPDFVTATGDTVEKSPFWSGDDHKNKHEKFPWFVSVINNYKVVDYQNRGTVP 308
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 208 ARSADPTFYFTALGDTVEVSPFCVDNSCPNATDVL--VQIDLNTVVDYGNRATSSQ 265

Qy 309 GKTRIFLDREBYTLISWEKHLKNMYSCLTLWKAFYNGIQTEHSGSYHFVANDITASTTS 368
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 266 HKKRIFAHTLDYSVSEAWKNSASVCSVMFWKSPQRAIQTEHDLTYHFIANEITAGFSTV 325

Qy 369 KEDMKEFTTTHCLNEEKABIEKYAKVNSTHSGYKGLKVFKTDGGLYLVMOPLIQNRL 428
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 326 KEPLANETSDYCNCLMTHINTTLEDKIARVNNTHTPNGTAIEYQTEGGMILVMOPLIAEL 385

Qy 429 LDAKNKLNNE-----TYSRRSRROAESTTDDPMEMTNGAGGEYSSENSITVAQVQYA 481
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 386 EEALEATTSPVTSAPTSSRSKRAIRSIRDV-----SAG-----SENNVFLSQIQYA 434

Qy 482 YDNURIRINILEDLSKAWCEQRAALVWNELSKINPTSMVSMIYNRPVSAKRIGDVIS 541
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 435 YDKLRQSIINNVLEBLAITWCREQVROTVMVVEIAKINPTSMVTAIYKGPVSRKALGDVIS 494

542 VSNCIIVDQTSVSLHKSRLLSASDEKCFSPRPVPTFKFMDNSTIYKQGLGVNNEILLTTT 601
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 495 VTECINVDQSSVSIHSLK--TENNDICYSRPPVPTFKFVNSSQLFKQGLGARNEILLSES 552

602 YLETCQENTYYPQAKTDMYIYKYNVHLKTVPLUSSITLDTFTALMFTLLENVDFKVIEL 661
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 553 LVENCHQNAETFTTAKNETHYFKNYVHVVELTPVNNISTLDTFLALNLTFTENIDFRAVEL 612

662 YTRDEKLSNVFDIETMFREYNVYQVSGRLKDLDDL--TNRNQFVDAFGSISMODDLGA 719
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 613 YSSGERKLANVFDLETMFREYNVYQVSGISGLRKD-FDNSQNRNDRDRIQDFSEILADLGS 671

720 VGTGVNAVSGVATLFSSTVGTGFINKPFGQMLMIIVVIGVFAIYFLLTKTKIYETA 779
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 672 IGKVIIVNAGAFSLFGGIVTGILNFIKNPLGCMFTFLGAVIILVILLVRETNMSQA 731

780 PIKMIYPEIDKLBKREKGEIAPISSEELERIVLAMHIHQNSHMETKTRKDKPKDSILTR 839
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 732 PIRMIYPDVEK-----SKSTVTPEPETIKQILLGMHNMQOEAYKKKEQRAARPSIFRQ 786

840 -AQNMLRKBSGYSNKLKNAESV 859
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 787 ASETPLURKSGYKQISTEDKI 807

RESULT 11
US-10-055-364-43
; Sequence 43, Application US/10055364
; GENERAL INFORMATION:
; APPLICANT: Patience, Clive
; TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use
; FILE REFERENCE: 61750-379
; CURRENT APPLICATION NUMBER: US/10/055,364
; CURRENT FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: US/09/612,204
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US/60/142,736
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: US/60/168,532
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 43
; LENGTH: 824
; TYPE: PRT
; ORGANISM: Eguine herpesvirus 2
; US-10-055-364-43

Query Match 42.7%; Score 1940; DB 26; Length 824;
Best Local Similarity 47.0%; Pred. No. 3.8e-173; Indels 64; Gaps 16;
Matches 395; Conservative 141; Mismatches 241;

Qy 11 VLALWLYQVALYSLSIAETGVTSPPTATWSTESPLTGHYTHDSHGERGNENRDE 70
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 10 VLCLWCV--AALLCQGVQAEVVA-----ETTP-----FATH-----RPEVVAE 46

Qy 71 EQNKNIYGPSPTFFRYVCSASGV-GDVFQDTHVCPDASDMVHSEGIILYKQNIIPFM 129
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 47 ENPANPF-----LPPFVCGASPTGGEIFRPPLESCPNTEDEKDHIEGIALIYKTNIVPV 101

Qy 130 FRYRKRVKVVTTSTVYNGIYSDITNQHTFYKIEPWETKMDTIYOCFNSRLNTGNNL 189
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 102 FNVKRYKRTMTSTIYKGSSEDAITNQHTSYAVPLYEVQMDHYIYQCSAVQVNEGCHV 161

Qy 190 LTYVDRDDINMTVFLQPDGVTPDVKRYGSOPELYLEPGWFGSYRRRTTVCNCLDMFPA 249
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 162 NTYYDRDGNWETAFLKPADGLTSSITRYQSOPEVYATPRNLLWSYTRTTRTVCNEVTEMSA 221

Qy 250 RSNPPDFVTATGDTVEKSPFWSGDDHKNKHEKFPWFVSVINNYKVVDYQNRGTVP 309
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 222 RSMKPPFEFFVTSYGDRTIEMSPFLKENGTEBPEKIKRPHSIQQLKKNYAVTKY----GVGLG 277

Qy 310 K-----TRIFLDREBYTLISWEKHLKNMYSCLTLWKAFYNGIQTEHSGSYHFVANDITAS 365
```


Db	278	QADNATREFFAI	FGDYSLSWKAT	TENSSYCDIL	LKNGFSNAI	QTOHNSLHFI	ANDITAGF	337		
Qy	366	TTSKEDMKFNT	TYICLNEEBE	IKAEIEKKYA	KVNSTHSGY	GDLYKFKD	GGLYLWVQLIQ	425		
Db	338	STPLEBEANFN	ETFKICWNNT	OEBIEQKKL	KEVEKTHRP	NGTAKVYKT	GNLYLWVQLIQ	397		
Qy	426	NRLLDAKNKN	-----	ETYSRRSR	ROAESTT	DDPMEMTG	NGAGGEYSS	472		
Db	398	IDLLDTHAKL	YNLTNATAS	PTSTPTSP	RRRRRTSS	-----	VSGGNGNGDN	450		
Qy	473	ITVAQQYAYD	NLRIRINNI	LEDLSKAC	RQHRRAAL	WNELSKIN	PTSVMSIYNRPVS	532		
Db	451	VAASQVOFAY	DNLRKRSIN	RVLGELS	RANCRQYR	ASLWYEL	SKINPTSVMSAIYGRPVS	510		
Qy	533	AKRIGDVISV	NCIVDQTS	VSLSLHKS	LLBSASDE	KCFSPRPV	TFKPMNDSTTYKGOOLGV	592		
Db	511	AKLIGDWSV	SCISVDQK	SVFVHKNM	K- VPGKED	CLCTRPV	VGVGFKPINGS	SELFAGOLGP	569	
Qy	593	NNEILITTY	LETCTOENT	EYFQAKT	DMYIYKNV	YEHLLK	TVPLSSITTT	LDTFIALNFTLLE	652	
Db	570	RNEIVLSTQ	VEVCQHS	CEHFQAGN	QMYKYKYV	YSTLNL	LDIPTLHTMITL	NLSLVE	629	
Qy	653	NVDKFVIELY	TRDEKRLS	NVEDIETM	FREYNYAQR	VSGLRKOL	LLD- LSTNRN	QOFVDFAG	711	
Db	630	NIDKFVIELY	SKTEKRLS	NVPDIETM	PREYNYV	QNLNGLRK	COLDDSDSIDHGR	DSFIQTLLG	689	
Qy	712	SLMDDLGA	VGQVNVNA	VSQVATL	PSSIVTG	FINIKNPF	PGGMLMI	IVVI- GVLFAY	IFYUT	770
Db	690	DIMQDLGTT	GKVVNVNA	VSQVSLF	GSIVSGV	ISFVKP	PGGMLLIV	LIAGV	VVYVLFMT	749
Qy	771	KKTIKYETAP	IKMYPED	DKLKEREG	KEIAPI	SEBLER	IVLAMIH	IQONSHWETK	K	830
Db	750	RRRSIY- S	APIRMLPG	VERAAQEP	CAH- - -	PVSEDQ	IRNILMG	HOHQFORQ	RABE	805
Qy	831	D	831							
Db	806	E	806							

RESULT 12

US-08-804-439-94	
Sequence 94, Application US/08804439	
GENERAL INFORMATION:	
APPLICANT: Rose, Timothy M.	
APPLICANT: Bosch, Marnix L.	
APPLICANT: Strand, Kurt	
TITLE OF INVENTION: GLYCOPROTEIN B OF THE RPHV/KSHV	
TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES	
NUMBER OF SEQUENCES: 113	
CORRESPONDENCE ADDRESS:	
ADDRESSEE: Morrison & Foerster	
STREET: 755 Page Mill Road	
CITY: Palo Alto	
STATE: CA	
COUNTRY: USA	
ZIP: 94304-1018	
COMPUTER READABLE FORM:	
MEDIUM TYPE: Floppy disk	
COMPUTER: IBM PC compatible	
OPERATING SYSTEM: PC-DOS/MS-DOS	
SOFTWARE: Patent In Release #1.0, Version #1.30	
CURRENT APPLICATION DATA:	
APPLICATION NUMBER: US/08/804,439	
FILING DATE:	
CLASSIFICATION: 424	
ATTORNEY/AGENT INFORMATION:	
NAME: Schiff, J. Michael	
REGISTRATION NUMBER: 40,253	
REFERENCE/DOCKET NUMBER: 29938-20002.20	
TELECOMMUNICATION INFORMATION:	
TELEPHONE: (415) 813-5600	
TELEFAX: (415) 494-0792	

Query Match	42.4%	Score	1927.5;	DB	14;	Length	829;
Best Local Similarity	44.2%;	Pred. No.	5.8e-172;				
Matches	391;	Conservative	145;	Mismatches	172;	Indels	103;
Gaps	15						

Qy	11	V L A L W L Y Q A L Y S L S I A E T G V T S P P N T A T W S T E S P L T G H Y G T H D S H G E R G N N E N R D S E	70
Db	10	L L R A W V V - - I I A I G T A V G E - N V T T P K A T T A K T P P - - - - -	42
Qy	71	E Q N K N I Y G S P S T - - - - - F P Y R V C S A G S G V D V F R Q T D H V C P D A S D M V H S E G I L	118
Db	43	- - - - - G P S T P T P P E N P P R A E A F R V C S A S A T G E L F R F N L E K T C P G T E D K T H Q E G I L	94
Qy	119	L I Y K Q N I I P M P F R V K Y R K V V T T S T V N G I Y S D S I T N Q H T F Y K S I E B W E T E K M D T I Y Q C F	178
Db	95	M V F K K N I V P H I F K V R R Y R K V A T S V Y V R G W T E A V T G K Q E V I R P V P Q Y E I N H M D T T T Q C F	154
Qy	179	N S L R L A T G G N L L T Y V D R D D I N M T V F L Q P V D G V T P D V K R Y G S O P E L Y L E P G W F M G S Y R R R T	238
Db	155	S S M R V N V N G L V N T T D R D F T N Q T V F L Q P V E G L T D N I Q R Y F S O P V L Y T T P G M F P G I Y R V R T	214
Qy	239	T V N C E L M D M F A R S N P P D F E V T A T G D T V E M S P F W S G E - - - - - D D H E N K M K H E K P F W S V I	292
Db	215	T V N C E I V D M I A R S A E P Y S Y F V T A L G D T V E V S P F C H N D S T C S V A E K T E N G L C A R - - - - - V L	269
Qy	293	N N Y K V D Y Q N R G T V P L G K T R I F I D R B E Y T L S W E K H L K N M S Y C P I T L W K A F Y N G I O T E H S G	352
Db	270	T N Y T M V D F A T R - - A P T E T R V A F D S G E Y T V S W K A E D P K S A C A L T L W K T F P R A I Q T T H E A	327

QY	353	SYHFVANDITASFTSKEDMKEFNNTYHCLNEIEKABIEKKYAKVNSTHYSKYGDLKYFKT	412
Db	328	SYHFVANDVATFTSPLESEVANFTGTVSCLEDEVIQKLTNTIKLSLTHVNTGSAQYKKT	387
QY	413	DGGLYLWQPLIQRLLDKAKNKNNETYS-----RRSRROQAESTTDPMMEMTGN	461
Db	388	EGGLFLLWQPLTPLSLVDEMRGLNGTTPAPPAITSTVSRYRVSVNTNEQATD-----	439
QY	462	GAGEYSSENSITVAQVQAYDNLIRINNILEDLSKAWCREQHRALVWNLSEKINPTS	521
Db	440	-----NLAAPQLQFAYDKLRASINKVLELSRAWCREQVDRDTYMYVLSKINPTS	489
QY	522	VMSMIYRNPVSXAKIEGIVSNICIVVDQTSVLSLHKSRLLSASDEKCFSPRPVTFKEWN	581
Db	490	VMTAIYGRPVSXAKPVGDAISVTDCAVDAQSVSIHKSRL--TSTPGICYSPRPVTFRFLN	547
QY	582	DSTIYKQGLGVNNEILTTTYLETQCENTEYFYQAKTDMYIKNYEHLKTVPLSSITLTD	641
Db	548	STTLFKQGLGRNEIILTDNQVEACKETCEHYFIASNVYKYDYVFKKINTSEISTLG	607
QY	642	TFIALNFTLNVDFKVIELYTRDEKRLS--NVFDIETMFREYNYAQRVSGRLKDL---L	697
Db	608	TFIALNLSFIENIDFRVIELYSRAEKKLSGSVFDIETMFREYNYTQRLAGLREDLNTI	667
QY	698	DLSTNRNQFVDAFGSLMDDLGAQCOTVNVAVSGVATLFFSSIVTGFNFINKNPFSGMLMI	757
Db	668	DL--NRDLRLDLSIADVADGDVGRVTVNVAVSSVITLFGSIVSGFINFKSPFGMLMIL	725
QY	758	VWIGVLPFAIYFLTKTKIYETAPIKMIYPIEIDKLEREGKSEIAPISEEELERIVLAMHI	817
Db	726	VIVAVVLIVPALNRRTNAIAQAIPRMYPIDIDKMQPSGGK-----VDQEIKNILAGMHQ	780
QY	818	HQONSHMETKTRDKPDKDSILTRAQNMLRKR--SGYSNLKNAESVE	860
Db	781	LQOEERRRLDEQORSAPSLFRASDGLKRRFRGVKYLENEEAQE	824
RESULT 16			
US-09-831-000-13			
; Sequence 13, Application US/09831000			
; GENERAL INFORMATION:			
; APPLICANT: Oregon Health Sciences University			
; TITLE OF INVENTION: Cloning of Rhadinovirus Genome and Methods for its Use			
; FILE REFERENCE: 53683			
; CURRENT APPLICATION NUMBER: US/09/831,000			
; CURRENT FILING DATE: 2001-05-02			
; PRIOR APPLICATION NUMBER: 60/107,507			
; PRIOR FILING DATE: 1998-11-06			
; PRIOR APPLICATION NUMBER: 60/109,409			
; PRIOR FILING DATE: 1998-11-20			
; NUMBER OF SEQ ID NOS: 179			
; SOFTWARE: Patent in Ver. 2.1			
; SEQ ID NO 13			
; TYPE: PRT			
; LENGTH: 829			
; ORGANISM: Macaca mulatta rhadinovirus 17577			
US-09-831-000-13			
Query Match 42.4%; Score 1927.5; DB 23; Length 829;			
Best Local Similarity 44.4%; Pred. No. 5.8e-172;			
Matches 388; Conservative 146; Mismatches 258; Indels 81; Gaps 14;			
QY	11	VLALWYLYQVALYSLSTAETGVSPNTATWSTESPLTGHYTHDSSHGRRGNENRDS	70
Db	10	LLRAWVV--IIAIGTAVGE--NVTTPKGATTTAKPTP-----	42
QY	71	EQKNKIYGSPT-----FPYRVCSAGVGQVDFRQTDHVCPSDAMVHSEGIL	118
Db	43	-----GSTPTPPNPRAAFKFRVCSAGATGELFRFNLEKTCPTGTEDKTHQEGIL	94
QY	119	LIYQNIIIPFMRVRYKRVVVTSTVYNGIYSDSITNQHTTFYKSIPEWETERKMDTYQCF	178
Db	119	LIYQNIIIPFMRVRYKRVVVTSTVYNGIYSDSITNQHTTFYKSIPEWETERKMDTYQCF	178
Db	95	MVFKNKIVPHIFKRRYRKVATSVTVYRGWTETAVTGQEVIRPVQYQVEINHMDDTYQCF	154
QY	179	NSLRNTGGLNLTYYVDRDDINMTVFLOPVDGVTDPVKRYGSOPELYLEPCWFWGSRRT	238
Db	155	SSMRVNVGIVNTYTDRTNQTQVFLQPVVEGTDNIQRYFSQPELVLYTTPGFWPGIYRVRT	214
QY	239	TVNCELMDMFASNPDPDFVTATGDTVEMSPFWSGE-----DDHENKHKHKEPWFVSVI	292
Db	215	TVNCELMDMFASNPDPDFVTATGDTVEMSPFWSGE-----DDHENKHKHKEPWFVSVI	269
QY	293	NNYKVVDYQNRGTVPGLKTRIFLDREBYTSLSEKHLKNMSYCPILTTLWKAIFYNGIOTESHG	352
Db	270	TNVTIVDFATR--OPTTETRVFADSGEYTVSKAEDPKSAVCALTTLWKTTPRAIQTHEA	327
QY	353	SYHFVANDITASFTSKEDMKEFNNTYHCLNEIEKABIEKKYAKVNSTHYSKYGDLKYFKT	412
Db	328	SYHFVANDVATFTSPLESEVANFTGTVSCLEDEVIQKLTNTIKLSLTHVNTGSAQYKKT	387
QY	413	DGGLYLWQPLIQRLLDKAKNKNNETYS-----RRSRROQAESTTDPMMEMTGN	461
Db	388	EGGLFLLWQPLTPLSLVDEMRGLNGTTPAPPAITSTVSRYRVSVNTNEQATD-----	439
QY	462	GAGEYSSENSITVAQVQAYDNLIRINNILEDLSKAWCREQHRALVWNLSEKINPTS	521
Db	440	-----NLAAPQLQFAYDKLRASINKVLELSRAWCREQVDRDTYMYVLSKINPTS	489
QY	522	VMSMIYRNPVSXAKIEGIVSNICIVVDQTSVLSLHKSRLLSASDEKCFSPRPVTFKEWN	581
Db	490	VMTAIYGRPVSXAKPVGDAISVTDCAVDAQSVSIHKSRL--TSTPGICYSPRPVTFRFLN	547
QY	582	DSTIYKQGLGVNNEILTTTYLETQCENTEYFYQAKTDMYIKNYEHLKTVPLSSITLTD	641
Db	548	STTLFKQGLGRNEIILTDNQVEACKETCEHYFIASNVYKYDYVFKKINTSEISTLG	607
QY	642	TFIALNFTLNVDFKVIELYTRDEKRLS--NVFDIETMFREYNYAQRVSGRLKDL---L	697
Db	608	TFIALNLSFIENIDFRVIELYSRAEKKLSGSVFDIETMFREYNYTQRLAGLREDLNTI	667
QY	698	DLSTNRNQFVDAFGSLMDDLGAQCOTVNVAVSGVATLFFSSIVTGFNFINKNPFSGMLMI	757
Db	668	DL--NRDLRLDLSIADVADGDVGRVTVNVAVSSVITLFGSIVSGFINFKSPFGMLMIL	725
QY	758	VWIGVLPFAIYFLTKTKIYETAPIKMIYPIEIDKLEREGKSEIAPISEEELERIVLAMHI	817
Db	726	VIVAVVLIVPALNRRTNAIAQAIPRMYPIDIDKMQPSGGK-----VDQEIKNILAGMHQ	780
QY	818	HQONSHMETKTRDKPDKDSILTRAQNMLRKR--SGYSNLKNAESVE	860
Db	781	LQOEERRRLDEQORSAPSLFRASDGLKRRFRGVKYLENEEAQE	824
RESULT 17			
US-10-055-364-37			
; Sequence 37, Application US/10055364			
; GENERAL INFORMATION:			
; APPLICANT: Patience Clive			
; TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use			
; FILE REFERENCE: 61750-379			
; CURRENT APPLICATION NUMBER: US/10/055,364			
; CURRENT FILING DATE: 2002-01-23			
; PRIOR APPLICATION NUMBER: US/09/612,204			
; PRIOR FILING DATE: 2000-07-07			
; PRIOR APPLICATION NUMBER: US/60/142,736			
; PRIOR FILING DATE: 1999-07-08			
; PRIOR APPLICATION NUMBER: US/60/168,532			
; PRIOR FILING DATE: 1999-12-02			
; NUMBER OF SEQ ID NOS: 55			
; SOFTWARE: Patent in version 3.0			
; SEQ ID NO 37			
; LENGTH: 823			
; TYPE: PRT			
; ORGANISM: Human herpesvirus 8			
US-10-055-364-37			

```
Query Match 42.2%; Score 1918; DB 26; Length 823;
Best Local Similarity 48.5%; Pred. No. 4.6e-171;
Matches 378; Conservative 134; Mismatches 233; Indels 34; Gaps 10;

QY 83 PPRVCSASGVDVRFQFQDTHVCPDASDMVHSEGLLIYKONIIPFMRVRYKRVKVTTS 142
DB 63 YOFVCSASITGELFRNLEQCPDQKQYHQEGILLVYKKNIVPHIEKVRRYKIATSV 122

QY 143 TVYNGIYSDSINQHTFYKSIPEWETEKMDTIYQCNSLRNTGGLNLTYYDRODDINMTV 202
DB 123 TVYRGLTSATINKYELPRPVLYEISHMDSTYQCFSSKVNUNGVNFTFDRDDVNTV 182

QY 203 FLQPDVGVTPDKRYGSOPELYLEPGWFGSYRRRTTNCMLDMFARNSPPDFVVTAT 262
DB 183 FLQPEGLTDNIQRYFSQFVIAEPGFWPGVIRVTVNCEIVDMIAESAEPYNYFVTSL 242

QY 263 GDTVEMSPFWGDEDDHENKMEKHPF-VSVINNYKVVDYQNRGTVPGLKTRIFLDREY 321
DB 243 GDTVEVSPFCYNSSCSTTPSNKGLSVQVVLNHTVWYSDRGTSPTPQNRIFVETGAY 302

QY 322 LSWKHLKNMSYCPULTLWAKAFYNGIOTBSHSGSYHFVANDITASFTTSKEDMKEFNTV 381
DB 303 LSWASEKTTAVCPALWKTTPRSIQTHDSHFHFVANEITATPTAPLTPVANFTDTYSC 362

QY 382 LNEEIKABIEKKYAKVNSTHSGYKDGKFKTDGGLYLVWQPLIQNRLLDAKNKLNET-- 439
DB 363 LSTDINTNLNASKAKALASTHPNGTVQYFHTTGGYLVWQPSALNTHAQDGSNTSS 422

QY 440 -----YRSRRQASSTTDPMMETGNGAGGYSSENSITVAQVAVDNLIR 488
DB 423 PPPSAPMTTASRRKRASASTA-----AAGGGGSDN-LSYTLQLOFAYDKLRDG 471

QY 489 INNILEDLKAWCREQHRALVWNLKINPTSVMSMIYNRPVSRAKRGDVISVNCITV 548
DB 472 INQVLEELUSRAWCREQVNDLMWYELSKINPTSVMTAIYGRVPSAKFVGDAISVTECIN 531

QY 549 DQTSVSLHSLRLLSASDEKCFSPVPVTFKPMNDSTIYKGLGVNNEILLTTTLYLETCOE 608
DB 532 DOSSVNIHSLR--TNSKDCVAPLVTFKELNSSLNLTGQLGARNEILLTNQVETCKD 589

QY 609 NTEYFOAKTDMYIYKNEYHEKTVPLSSITLDTFIALNFTLLENVDPKVIELYTRDEKR 668
DB 590 TCEHYFTRNETLVYKDYAVLRTINTDITLNTFIALNLSFIQIDFKALIELYSSAEKR 649

QY 669 L-SNVFDIETMFREYNYAQRVSLGRKDLDD--LSTNRNQFVDAFGLSMDLLGAVGQTVN 726
DB 650 LASSVFDLETMFREYNYTHRLAGLREDLNDTIDMKNKERFVRDLSEIVADLGGIGKTVN 709

QY 727 AVSGVATLFSSTVTFINFIKNPFGGMLMIWIGLVFAIYFLTKTKIYETAPIKMIYP 786
DB 710 VASSVVTLCGSLVTGFINFIKHPGLGMLMIIVIAIILIIIPMLSRNTTIAQAPVKMIYP 769

QY 787 BIDKLEREGKSEIAPISEEELERIVLAWHIHQNSHMETKTRDKPKDSILTRAQNMLR 845
DB 770 DVDR---RAPPSGGAP-TREEIKNLLGWHQLOQBRQKADDLKKS-TPSVFQRTANGLR 823
```

RESULT 18

```
US-10-055-364-42
; Sequence 42, Application US/10055364
; GENERAL INFORMATION:
; APPLICANT: Patience, Clive
; TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use
; FILE REFERENCE: 61750-379
; CURRENT APPLICATION NUMBER: US/10/055,364
; CURRENT FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: US/09/612,204
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US/60/142,736
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: US/60/168,532
; PRIOR FILING DATE: 1999-12-02
```

```
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 42
; LENGTH: 792
; TYPE: PRT
; ORGANISM: Herpesvirus saimiri
US-10-055-364-42
```

Query Match 42.1%; Score 1916; DB 26; Length 792;

```
Best Local Similarity 46.9%; Pred. No. 6.6e-171;
Matches 366; Conservative 155; Mismatches 228; Indels 32; Gaps 9;
```

```
QY 71 EQNKNIYGSPTPYRVCSASGVDVRFQFQDTHVCPDASDMVHSEGLLIYKONIIPMF 130
DB 28 EKNKTQAIYQBYFYRVCSASTTGCFLFRDLDRCTPSTEDKVKHEGILLVYKKNIVPIF 87

QY 131 RVRKYRKVVTTSTVYNGIYSD--SITQHTFYKSIPEWETEKMDTIYQCNSLRNLTCGN 188
DB 88 KVRREYKITTTSVRIFNGWTRGVAITNKWLSRAVPKYEIDIMDKTYQCHNCMQIEVNGM 147

QY 189 LLTYVDRDDINMTVFLQPDVGTDPVKRYGSOPELYLEPGWFGSYRRRTTNCMLDMF 248
DB 148 LNSYDGRGNKTVDLAPVDGLTGCAITRYISQPKVADPGLWGTYTRITRVNCEIVDMF 207

QY 249 ARSNPPDFFTATGDTVEMSPFWGDEDDHENKMEKHPFVSVINNYKVVDYQNRGTVP 308
DB 208 ARSADPTVYFVTALGDTVEVSPFCVDVNSCPNATDVLS--VQIDLNHTVVDYGNRATSQ 265

QY 309 GKTRIFLDREYTLWSWEKHLKNMSYCPULTLWAKAFYNGIOTBSHSGSYHFVANDITASFTTS 368
DB 266 HKKRFIFATLTDYSVSEAVNKSASVCSVMVFWKSFQRAIOTEDHDLTYHFIAEITAGFSTV 325

QY 369 KEDMKENTTYVHCLEEBIEKAEIEKKYAKVNSTHSGYKDGKFKTDGGLYLVWQPLIQNR 428
DB 326 KEPLANTSDYNCMLTINTTLEDKIAKRVNHTTPNGTAETIYQTEGGMLVWQPLIAIEL 385

QY 429 LDANKKLNE-----TYSRRSRQASSTTDPMMETGNGAGGYSSENSITVAQVQYA 481
DB 386 EAMLEATTSPVTPSAPTSSRSKRAIRSIRDV-----SAG-----SENNVFLSQIOYA 434

QY 482 YDNLRIKINILEDLKAWCREQHRALVWNLKINPTSVMSMIYNRPVSRAKRGDVIS 541
DB 435 YDKLRQSIINNLEELAITWCREQVQRQTMVWVEIAKINPTSVMTAIYKGPVSRKALGDVIS 494

QY 542 VSNICIVDQTSVLSHSLRLLSASDEKCFSPVPVTFKPMNDSTIYKGLGVNNEILLTTT 601
DB 495 VTECINVDQSSVSIHSLK--TENNDICYSRPPVTFKPVNSSQLFKGQLGARNEILLSES 552

QY 602 YLETQCENTEYFOAKTDMYIYKNEYHEKTVPLSSITLDTFIALNFTLLENVDPKVIEL 661
DB 553 LVENCHQNAETFFTAKNETHFKNYVHVELPVNNISTLOTFLALNLTFFIENIDPKAVEL 612

QY 662 YTRDEKLSNVFDIETMFREYNYAQRVSLGRKDLDDLS--TNRNQFVDAFGLSMDLLGA 719
DB 613 YSSGERLANVFDLETMFREYNYAQSISGLRKD-FDNSQRNRRDRIIQDFSEILADLGS 671

QY 720 VGQTVNAVSGVATLFSSTVTFINFIKNPFGGMLMIIVIGLVFAIYFLTKTKIYETA 779
DB 672 IGKVINNVASGAFSLFGGIVTGILNFIKNPLGGMFTFLIGAVIILVILLVRRNNNSQA 731

QY 780 PIKMIYPEIDKLEREGKSEIAPISEEELERIVLAWHIHQNSHMETKTRDKPKDSILTR 839
DB 732 PIRMIYFDVEK-----SKSVTPMEPETIKQILLGMHMQOEAYKKKEQBAARPSIFRQ 786

QY 840 A 840
DB 787 A 787
```

RESULT 19

```
US-10-055-364-38
; Sequence 38, Application US/10055364
; GENERAL INFORMATION:
```

```
; APPLICANT: Patience, Clive
; TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use
; FILE REFERENCE: 61750-379
; CURRENT APPLICATION NUMBER: US/10/055,364
; CURRENT FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: US/09/612,204
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US/60/142,736
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: US/60/168,532
; PRIOR FILING DATE: 1999-12-02
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 38
; LENGTH: 808
; TYPE: PRT
; ORGANISM: Rhesus monkey rhadinovirus
US-10-055-364-38

Query Match 41.8%; Score 1902; DB 26; Length 808;
Best Local Similarity 44.2%; Pred. No. 1.5e-169;
Matches 384; Conservative 144; Mismatches 240; Indels 102; Gaps 14;

Qy 11 VLALWLYQVALYSLSIAETGVTSPNTATWSTESPTLGHYTHDSSHGERGNENRDSE 70
Db 10 LLRAWV--IAIGTAVGE-NVITPKGATTAKTP-----42
Qy 71 EQKNKIYGSPT-----FPYRVCASGVGVFRFQTDHVCPCDASDMVHSEGL 118
Db 43 -----GRSTPPENPPRAEAFKFRVCASATGELFRNLEKTCPTGTEKTHQEGIL 94
Qy 119 LIYKQNLIPMERVKRYKVVTTSTVNGIYSDISITNQHTFYKSIETPEWETEKMDTIYQCF 178
Db 95 MVFKNLIVPHIFKRRYRKVAITSVTVRGWTTAVTGQEVIRPQVQIEINHMDTIYQCF 154
Qy 179 NSLRNLNTGNNLLTVYVDRDDINMTVFLQPDVGVTPDVKRYGSPQLYLEPGWFGWSYRRRT 238
Db 155 SSMEVNVNGIVNTYDRDFTNQTVFLQPDVGVTPDVKRYGSPQLYLEPGWFGVIYVRT 214
Qy 239 TVNCELMDMFARSNPPDFVFTATGDTVEMSPFWSGE-----DDHKNKHEKPFWFSVI 292
Db 215 TVNCEIVDMIAESAEPYSYFTALGDTVEVSPFCDAENSCPNAS-----VL 269
Qy 293 NNYKVVQYQNGTVPLCKTRIFLDREBYTILSWKHLKNMSYCLTLWKAFYNGIOTHEHSG 352
Db 270 TNYTMDVFAIR--APTETRVFDSGTYTVSWKAEADPKSAVCALTMLKTPPRAIQTTHA 327
Qy 353 SYHFVANDITASFTTSKEDMKFNTHYHCLNEEIKAEIEKKYAKVNSTHSGYDGLKYFKT 412
Db 328 SYHFVANDVATFTSPVSEVANFTGYSCLDDEVILQKTLNDTIKLSDTHVNTGSAQYKT 387
Qy 413 DGGLYLVWQPLIQNRLDANKLNNETYS-----RRSRROAESTTDPMMWNTGN 461
Db 388 EGGFLFWLQPLTPLSLVDENRGLTTPAPPATTSTVSRVRRSVNTNEQAT-----439
Qy 462 GAGEYSSSENSITVAQVQAYDNLRIINNILEDLKAWCREQHRALVWNLKINPTS 521
Db 440 -----NLAAPOLOQFAYDKURASINKVLEBSRAWCREQVDTYMMYELSKINPTS 489
Qy 522 VMSMIYNRPSAKRIGDVISVNSCIVVDQTSVSLHKSRLLSASDEKCFRPPVTFKFMN 581
Db 490 VMTAIYGRPSAKRIGDVISVNSCIVVDQTSVSLHKSRLLSASDEKCFRPPVTFKFMN 547
Qy 582 DSTIYKQGLGNNHLLTITVYTCQENTYFOAKTDMYIKNVEHLKTVLPSLSSITLD 641
Db 548 STTLFKQGLGRNEIILTDNQVEACKETCEHYFIASNVTVYKYDYVFKKINTSEISTLG 607
Qy 642 TETALNFTLENVDKVIELYTRDEKLS-NVFDIETMREYNKYAORVSLRKL- --L 697
Db 608 TETALNUSFTENDFRVIELYSRAEKLKSGSVFDIETMREYNKYAORVSLRKL- --L 667
Qy 698 DLSTNRNQVDAFGSLMDDLGAVGQTVVNAVSGVATLFFSSIVTGFIFNFKNPFEGMLMII 757
Db 698 LGDVISVTECINVDQTSVSIHSLK--TTNNDVCYSRPPVTFKFMNSSLQFKQGLGARNE 545

668 DL--NRDLRLADLSEIVADIGDVGRVWVNVASSVITLFGSIVSGFINFIKSPFGGMLMIL 725
Qy 758 VVIGVLFAYFLTKTKIYETAPIKMIYPEIDKLEREGKSEIAPISEELEERIVLAMHI 817
Db 726 VIVAVVLIIVFALNRRTNAIAQAPIRMIYDPIDKMQPSGGK-----VDQEQIKNILAGMHQ 780
Qy 818 HQONSHMETKTRKDKPKDSILTIAQNMLR 845
Db 781 LQOEERRRLDEQORSAPSLFRRASDGLK 808

RESULT 20
US-10-055-364-41
; Sequence 41, Application US/10055364
; GENERAL INFORMATION:
; APPLICANT: Patience, Clive
; TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use
; FILE REFERENCE: 61750-379
; CURRENT APPLICATION NUMBER: US/10/055,364
; CURRENT FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: US/09/612,204
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US/60/142,736
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: US/60/168,532
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 41
; LENGTH: 791
; TYPE: PRT
; ORGANISM: Ateline herpesvirus 3
US-10-055-364-41

Query Match 41.8%; Score 1898.5; DB 26; Length 791;
Best Local Similarity 46.6%; Pred. No. 3e-169;
Matches 367; Conservative 144; Mismatches 235; Indels 41; Gaps 10;

Qy 69 SEBQKNKIYGSPTFPYRVCASGVGVFRFQTDHVCPCDASDMVHSEGLIYKQNIIPF 128
Db 26 SDENGTGTPAIEKEYFYKRYVCSATGELFRNLDRACTEDKVRHREGILLVYKKNIVPH 85
Qy 129 MPVRVRYKRVVTTSTVNGIYSD--SITNOHTFYKSIETPEWETEKMDTIYQCFNSRLNTG 186
Db 86 IFKVRKYKIAISVRIFNGWSREGVAITKNWLSRAVPKVEINLMDKNVQCHNCMQIEVN 145
Qy 187 GNLLTVYVDRDDINMTVFLQPDVGVTPDVKRYGSPQLYLEPGWFGWSYRRRTTVNCELMD 246
Db 146 GLINSYCDRGNKNTVDLKPVDGLTGATIRYVSQPKIFADAGWLGWGTYKTRTTVNCIVE 205
Qy 247 MPARSNPPDFVFTATGDTVEMSPFWSGEEDDHENKHEKPFWFSVIN-----NYKVVQYQ 301
Db 206 MFARSADPTTYFTALGDTVEVSPFCDAENSCPNAS-----DVLSSQVDFNHTVVDYG 258
Qy 302 NRGCTVPLGKTRIFLDREBYTILSWKHLKNMSYCLTLWKAFYNGIOTHEHSGSYHFVANDI 361
Db 259 NRATSOQHGKRIFAHTLDYSVWEAINKTSVCSMVFWKGFQRAIQTEDHSTTHFANEI 318
Qy 362 TASFTTSKEDMKFNTHYHCLNEEIKAEIEKKYAKVNSTHSGYDGLKYFKTDCGLYLWQ 421
Db 319 TAGFSTSKETLASFSSEYSLMSDINSTLTDKIGRVNTHVPNGTAQYKPKTEGGMILVWQ 378
Qy 422 PLIQNRLDANKLN- -----ETYSRRRROAESTTDPMMWNTGNGAGEYSSSENSITV 475
Db 379 PLTAIELEEAITEATTVSPTPSTAHLTSRRTRGRKRDV-----SAG-----SENSVLL 427
Qy 476 AQVQAYDNLRIINNILEDLKAWCREQHRALVWNLKINPTSVMMSIYNRPSYAKR 535
Db 428 AQIQAYDKLROSINNVLLEELAITWCREQVQRTMIWEIAKINPTSVMTAIYGRPSAKA 487
Qy 536 IGDVISVNSCIVVDQTSVSLHKSRLLSASDEKCFRPPVTFKFMNDSTIYKQGLGNNNE 595
Db 488 LGDVISVTECINVDQTSVSIHSLK--TTNNDVCYSRPPVTFKFMNSSLQFKQGLGARNE 545
```


Db 457 ATVOIQAYDSURRQINRMGLDLARAWCLEQKQKQNNVRLTKINPTTVMSSIIYKAVAA 516
QY 534 KRIGDIVISVNCIVVDQTSVSLHSLRLLSASDEKCFGRPPVTFKFMNDSTIYKQGLGVN 593
Db 517 KRLGDVIVSQCPVNOATVTLRKSMR-VPGSETWCYSRPLVSPFINDTKTYEQGLGTD 575
QY 594 NEILLTYYLETCQNTYFPOAKTDMYIKNYEHLKTVPLSSITTLDTFIALNFTLLEN 653
Db 576 NEIFLTKMTEVCQATSOYFQSGNEIHVYNDYHHFKTIELDGIATLQTFISLNTSLIEN 635
QY 654 VDFKVIETYRDEKRLSNVFDIETMFREYNYAQRVSGRLKDLDD-LSTNRNQFVDAFGS 712
Db 636 IDFASLELYSRDEQRASNVFDLEGIFREYNFOAQNIAGRLKDLDDNAVSGNRNQFVDGLGE 695
QY 713 LMDDLGAVGQTVVNAVSGVATLFSISIVTGFINFKNPFPGMLMIIVIGVILFAIYFLTKK 772
Db 696 LMDSLGSGVQSITNLVSTVGGFLSSLVSGFISFFKNPFGGMLIILVAVGVILVLSLTR 755
QY 773 TKIYETAPIKMIYPIIDKLKEREKSE---IAPISEEBELERIVLAMHIHQONSHMETKTR 829
Db 756 TRQMSQCPVQMLYPCIDELAQHASGEGPGINPIKTELQAIMLA--LHEQNOEQKRAAQ 813
QY 830 KDPKDSILTRAQNMRLKR 847
Db 814 RAAGPSVASRALQAARDR 831

RESULT 24

US-09-301-390-18
; Sequence 18, Application US/09301390
; GENERAL INFORMATION:
; APPLICANT: Rose, Timothy M.
; APPLICANT: Bosch, Marnix L.
; APPLICANT: Strand, Kurt
; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RFHV/KSHV
; TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/301,390
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/720,229
; FILING DATE: 26-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Schiff, J. Michael
; REGISTRATION NUMBER: 40,253
; REFERENCE/DOCKET NUMBER: 29938-20002.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 857 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-301-390-18

Query Match 40.7%; Score 1848.5; DB 17; Length 857;
Best Local Similarity 46.1%; Pred. No. 1.9e-164;
Matches 368; Conservative 148; Mismatches 241; Indels 41; Gaps 12;
QY 81 STFPYRUCSAGVGDVFRFQTDHVCDD-ASDMVISEGILLIYKONIIPFMRVKRYKVV 139
Db 44 TSFFPRVCELSHGLDFRFSDDIQCPSPGTRENHTEGLLMVFKONIIPYSFKVRSYTKIV 103
QY 140 TTSVYNGIYSDSTINOHFTFYKTEPWETEMDITTYOCFNSLRNLTGNGLLTYVDRDDIN 199
Db 104 TNLIIYNGVADSVNHEEKFSVDSYETDMDIYQCYNAVKMTKGLTRVYVDRGVN 163
QY 200 MTVELQPVGDVTPDKRYGSOPELYLEPGFWGYSRRRTTVNCELMDFAKSNPFPDFV 259
Db 164 ITVNLKPTGGLANGVRRYASQTELYDAPGMLIWTYRTRTTVNCILITDMAKSNPFPDFV 223
QY 260 TATGDTVEMSPFSGEDDHENKMEKFWFVSVINNYKVVDYQNRGTVPGLKTRIFLDREE 319
Db 224 TTTGQTVEMSPFYDGK--NKETPHERADSFHVRNYKIVDYDNRGTNPQGERAPFLDKGT 281
QY 320 YTLSEWEXHLKN-MSYCPCLTLWKAFYNGIQTEHSGSYHFVANDITASFTTSKEDMKFNTT 378
Db 282 YTLSEW--LENRTAYCPLOHMQTFDSTIATETGSIHFVTDGTSSTVNTTVGIELPDA 339
QY 379 YHCLNEEIKABIEKKYAKVNSTHSGYGD-LKYFKTDGGLYLVWQPLIQNRLLDKRN--KL 435
Db 340 PKCIEEQVNTMHEKYEAVQDRYTKGOEATYFITSGLLLAWLPLTPRSLATVKNLTTEL 399
QY 436 NNETYSS-----RRSRRQAESTTDDPMEMWTGNGAGEYSSSENSI 473
Db 400 TPTTSSPPSPSPAPSAARGSTPAALVLRRRRRDAGNATTP---VPPTAPGKSLGTLLNP 456
QY 474 TVAQVOYADNLRIRINNILEDLSKACRQHRALVWNLKINPTSVMSIYNRPVSA 533
Db 457 ATVOIQAYDSLRQINRMGLDLARAWCLEQKQKQNNVRLTKINPTTVMSSIIYKAVAA 516
QY 534 KRIGDIVISVNCIVVDQTSVSLHSLRLLSASDEKCFGRPPVTFKFMNDSTIYKQGLGVN 593
Db 517 KRLGDVIVSQCPVNOATVTLRKSMR-VPGSETWCYSRPLVSPFINDTKTYEQGLGTD 575
QY 594 NEILLTYYLETCQNTYFPOAKTDMYIKNYEHLKTVPLSSITTLDTFIALNFTLLEN 653
Db 576 NEIFLTKMTEVCQATSOYFQSGNEIHVYNDYHHFKTIELDGIATLQTFISLNTSLIEN 635
QY 654 VDFKVIETYRDEKRLSNVFDIETMFREYNYAQRVSGRLKDLDD-LSTNRNQFVDAFGS 712
Db 636 IDFASLELYSRDEQRASNVFDLEGIFREYNFOAQNIAGRLKDLDDNAVSGNRNQFVDGLGE 695
QY 713 LMDDLGAVGQTVVNAVSGVATLFSISIVTGFINFKNPFPGMLMIIVIGVILFAIYFLTKK 772
Db 696 LMDSLGSGVQSITNLVSTVGGFLSSLVSGFISFFKNPFGGMLIILVAVGVILVLSLTR 755
QY 773 TKIYETAPIKMIYPIIDKLKEREKSE---IAPISEEBELERIVLAMHIHQONSHMETKTR 829
Db 756 TRQMSQCPVQMLYPCIDELAQHASGEGPGINPIKTELQAIMLA--LHEQNOEQKRAAQ 813
QY 830 KDPKDSILTRAQNMRLKR 847
Db 814 RAAGPSVASRALQAARDR 831

RESULT 25

US-09-338-326-18
; Sequence 18, Application US/09338326
; GENERAL INFORMATION:
; APPLICANT: Rose, Timothy M.
; APPLICANT: Bosch, Marnix L.
; APPLICANT: Strand, Kurt
; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RFHV/KSHV
; TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster

```
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/338,326
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/720,229
FILING DATE: 26-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Schiff, J. Michael
REGISTRATION NUMBER: 40,253
REFERENCE/DOCKET NUMBER: 29938-20002.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 813-5600
TELEFAX: (415) 494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 857 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-338-326-18

Query Match 40.7%; Score 1848.5; DB 17; Length 857;
Best Local Similarity 46.1%; Pred. No. 1.9e-164;
Matches 368; Conservative 148; Mismatches 241; Indels 41; Gaps 12;

QY 81 STFPYRVCASGVGDVFRFQTDHVCDD-ASDMVHSEGLLIYKQNIIPFMFRVKYRKV 139
DB 44 TSFPFRVCELSHGDLFRFSDIQCPFGTRENHTEGLLMVFKDNIIPYSFKVRSYTKIV 103
QY 140 TTSVYNGIYSDSITNOHTFYKSIPEWTEKMDTIYOCFNSLRNTGNNLLTYVDRDIN 199
DB 104 TNLIYNGWYADSVTNHREKFSVDSYETQMDTIYQYNAVXMTKDGLTRVIVDRDGVN 163
QY 200 MTVFLQPDVGTDPVKRYGSQLPELYLEPGFWGYSRRRTTVNCELMDMFARSNPPDFV 259
DB 164 ITVNLKPTGGLANGVRRYASQTELYDAPGMLIWTYRTRTTVNCILITDMAKSNPPDFV 223
QY 260 TATGDTVEMSGPFWGSDHDKMKHEKFPVSVINNYKVVDYQNRGTVPGLKTRIFLDREE 319
DB 224 TTTGQTVMSPFPYDGK--NKETFERADSFHVRTNYKIVDYNRGTNPQGERAFLDKGT 281
QY 320 YTLISWEKHLKN--MSYCPLTLWKAFYNGIQTEHSGSYHFVANDITASFTTSKEDMKERNIT 378
DB 282 YTLISWK--LENRTAYCPLQHWQTFDSTIATETGKSIHFVTDGTSFVNTTVGIELPDA 339
QY 379 YHCLNEEIKAEIKKAKVNSTHSGYGD-LKYFKTDGGLYLVNQPLFQNRLLDAKN--KL 435
DB 340 FKCIIEQVNTMKHEKYEAVQDRTYKQGEAITFYITSGGLLLAWLPLTPRSLATVKNLT 399
QY 436 NNETY-----RRSRQAESTTDPMMEMTNGAGGEYSSENSI 473
DB 400 TPTSSPFPSPPPAPSAARGSTPAALVRLRRRRDAGNATTP--VPPTAPKSLGLTLNPP 456
QY 474 TVAQVQAYNLRIRINILEDLSKAWCRQHRALVWNLKSKINPTSVMSIYNRPVSA 533
DB 457 ATVOIQFAYSLRQIRNMLGDLARAWCLQKQWVLRRLTKINPTTVMSIYKAVAA 516
QY 534 KRIGDVISVNCIVVDQTSVLSLHKSRLLSASDEKCFSPRPPTFKFMDNSTIYKQGLGVN 593
DB 517 KRLGDVISVQCPVQVQATVTLRSMR--VPGSETMCSYRPLVSPFSFINDTKTYEGQLGTD 575

QY 594 NEILLTTTLETCOENTYXXFOAKTDMYIKYKNEHLKTVPLSSITTTLDITFIALNFTLEN 653
DB 576 NEIFLTKMTEVCQATSQYIFQSGNEIHYNDVHHFKTIELDGIATQIFISUNSLIEN 635
QY 654 VDFKVIELYTRDEKRLSNVDFDIETMREYNYAQRVSGLRKDLID--LSTNRNDFVDAFGS 712
DB 636 IDFASLELYSRDEQRASNVDFLEGIFREYNFQAQNIAGLRKOLDNAVSNCRNQFVDGLGE 695
QY 713 LMDLGAQGVTVVNAVSGVATLFPSSIVTGFINFIPKPNPFGGMLMIIVVIGVLFAYFLTKK 772
DB 696 LMDSLGSGOSITNLVSTVGLFSSLVSGFISPFKNPFGGMLILVLVAGVILVISLTR 755
QY 773 TKIYETAPIKMIYPEIDKLKEREGKSE---IAPISEELERIVLAMIHOONSHMETKTR 829
DB 756 TROMSQOPVOMLYPGIDELAQOQHASGEGPGINPISKTELQAIMLA--LHEQNEQKRAAQ 813
QY 830 KDPKDSILTRAQNMRLKR 847
DB 814 RAAGPSVASRALQAARDR 831

RESULT 26
US-10-055-364-46
Sequence 46, Application US/10055364
GENERAL INFORMATION:
APPLICANT: Patience, Clive
TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use
FILE REFERENCE: 61750-379
CURRENT APPLICATION NUMBER: US/10/055,364
CURRENT FILING DATE: 2002-01-23
PRIOR APPLICATION NUMBER: US/09/612,204
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: US/60/142,736
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: US/60/168,532
PRIOR FILING DATE: 1999-12-02
NUMBER OF SEQ ID NOS: 55
SOFTWARE: Patentin version 3.0
SEQ ID NO 46
LENGTH: 829
TYPE: PRT
ORGANISM: Epstein-Barr virus
US-10-055-364-46

Query Match 40.6%; Score 1844.5; DB 26; Length 829;
Best Local Similarity 46.1%; Pred. No. 4.2e-164;
Matches 367; Conservative 148; Mismatches 240; Indels 41; Gaps 12;

QY 81 STFPYRVCASGVGDVFRFQTDHVCDD-ASDMVHSEGLLIYKQNIIPFMFRVKYRKV 139
DB 44 TSFPFRVCELSHGDLFRFSDIQCPFGTRENHTEGLLMVFKDNIIPYSFKVRSYTKIV 103
QY 140 TTSVYNGIYSDSITNOHTFYKSIPEWTEKMDTIYOCFNSLRNTGNNLLTYVDRDIN 199
DB 104 TNLIYNGWYADSVTNHREKFSVDSYETQMDTIYQYNAVXMTKDGLTRVIVDRDGVN 163
QY 200 MTVFLQPDVGTDPVKRYGSQLPELYLEPGFWGYSRRRTTVNCELMDMFARSNPPDFV 259
DB 164 ITVNLKPTGGLANGVRRYASQTELYDAPGMLIWTYRTRTTVNCILITDMAKSNPPDFV 223
QY 260 TATGDTVEMSGPFWGSDHDKMKHEKFPVSVINNYKVVDYQNRGTVPGLKTRIFLDREE 319
DB 224 TTTGQTVMSPFPYDGK--NKETFERADSFHVRTNYKIVDYNRGTNPQGERAFLDKGT 281
QY 320 YTLISWEKHLKN--MSYCPLTLWKAFYNGIQTEHSGSYHFVANDITASFTTSKEDMKERNIT 378
DB 282 YTLISWK--LENRTAYCPLQHWQTFDSTIATETGKSIHFVTDGTSFVNTTVGIELPDA 339
QY 379 YHCLNEEIKAEIKKAKVNSTHSGYGD-LKYFKTDGGLYLVNQPLFQNRLLDAKN--KL 435
DB 340 FKCIIEQVNTMKHEKYEAVQDRTYKQGEAITFYITSGGLLLAWLPLTPRSLATVKNLT 399
```

```

QY 436 NNETY-----RRRQRAESTTDPMMVTGAGGGEYSSENSI 473
Db 400 TTPTSSPPSPAPSAAGSTPAAVLRRRRDAGNATP---VPPTAPKSLGTLLNP 456
QY 474 TVAQVAYDNLRIIRINILEDLKAWCEQHRALVWNELSKINPTSVMSIYNRPVSA 533
Db 457 ATVQIQFAYDSURRQINRMGLDLARAWCLEQRQNMVRLTKINPTTVMSSIIYKAVAA 516
QY 534 KRIGDVISVNCIVVDQTSVLHKSRLLSASDEKCFGRPPVTFKPFMDSTIYKQGLGVN 593
Db 517 KRLGDVISVQCVPVQAATVLRKSMR--VPGSETWCYSEPLVSFFINDTKTYEGQLGTD 575
QY 594 NEILLTTTLETCQENTYFFQAKTDMYIKYKNEHLKTVPLSSITTLDTFIALNFTLLEN 653
Db 576 NEIFLTKKMTVCQATSOYYFQSGNEIHVYNDYHHFKTIELDGIATLQTFISLNTSLIEN 634
QY 654 VDFKVIELTRDEKRLSNVFDIETMFREYNYAQRVSGRLKDLDD--LSTNRNQFVDARGS 712
Db 636 IDFASLELYSDEQRASNVFDLEGIFREYNFOAQNIAGRLKDLDAVNSGRNQFVDGLGE 694
QY 713 LMDDLGAVGQTVVNAVSGVATLFSISIVTGFINFKNPFGGMLMIIVIGVLFAYFLTKK 772
Db 696 LMDSLGSGVQSITNLVSTVGGFLFSSLVSGFISFFKNPFGGMLIIVLVAGVILVISLRR 754
QY 773 TKIYETAPIKMIYPEIDKLEREGKSE---IAPISEEBELERIVLAWHIHQNSHMETKTR 829
Db 756 TRMSQQPQVMLYPGIDELAQOAHASGEGPGINPISKTELQAIMLA--LHEQNOEQKRAAQ 812
QY 830 KDPKDSILTTRAQNMLR 847
Db 814 RAAGPSVASRALQAADR 830

```

RESULT 27

```

US-08-471-913-103
; Sequence 103, Application US/08471913
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING EPSTEIN-BARR VIRUS
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471,913
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-030
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 103:

```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 856 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-471-913-103

```

Query Match 40.4%; Score 1839; DB 8; Length 856;

Best Local Similarity 46.1%; Pred. No. 1.5e-163; Matches 368; Conservative 148; Mismatches 240; Indels 42; Gaps 13;

```

QY 81 STFPYRVCSAGSGVDVFRFQTDHYVCPD--ASDMVHSEGIILYKQNIIPMFVRVRYKRVV 139
Db 44 TSFPRVCELSSHGDLPFRSSDIQCPSPGFTRENHTEGLLMVFKDNIIPYFKVRSYTKIV 103
QY 140 TTSVYNGIYSDSTINQHTFYKSIPEWETEKMDTIYQCFNSRLNLTGNNLLTYVDRDDIN 199
Db 104 TNLIIYNGWYADSVNRHEEKFSVDSYETQMDTIYQCYNAVVKMTKQGLTRVYVDRDGVN 163
QY 200 MTFVLQPDVGVTPDVKRYGSOPELYLEPGWFWGVRRTTVNCELMDMFARSNPPDFEV 259
Db 164 ITVNLKPTGGLANGVRRYVASQTELYDAPGWLITWYRTRTVNCLITDMAKSNPPDFEV 223
QY 260 TATGDTVEMSPFSGEDDHENKMHKPFVSVINNYKVVDYQNRGTVPGLKTRIFLDREE 319
Db 224 TTTGQTVEMSPFYDGK--NKETPHERADS PHVRTNYKIVDYDNRGTNPQGERRAFLDKGT 281
QY 320 YTLSEWELKKN--MSYCPDLTLWAFYNGIOTHSYSHYFVANDITASFITTSKEDMKFNTT 378
Db 282 YTLSEWELKKN--MSYCPDLTLWAFYNGIOTHSYSHYFVANDITASFITTSKEDMKFNTT 378
QY 379 YHCLNEEKAEIEKKYAKVNSTHSGYGD--LKYFTDGLGLVLMQPLIQNRLLDKN--KL 435
Db 340 FKCEEQVN--KTHEKYEAVQDRYTKGOEAITFYITSGLLLAWLPLTPRSLATVKNLTEL 398
QY 436 NNETY-----RRRQRAESTTDPMMVTGAGGGEYSSENSI 473
Db 399 TTPTSSPPSPAPSAAGSTPAAVLRRRRDAGNATP---VPPTAPKSLGTLLNP 455
QY 474 TVAQVAYDNLRIIRINILEDLKAWCEQHRALVWNELSKINPTSVMSIYNRPVSA 533
Db 456 ATVQIQFAYDSURRQINRMGLDLARAWCLEQRQNMVRLTKINPTTVMSSIIYKAVAA 515
QY 534 KRIGDVISVNCIVVDQTSVLHKSRLLSASDEKCFGRPPVTFKPFMDSTIYKQGLGVN 593
Db 516 KRLGDVISVQCVPVQAATVLRKSMR--VPGSETWCYSEPLVSFFINDTKTYEGQLGTD 574
QY 594 NEILLTTTLETCQENTYFFQAKTDMYIKYKNEHLKTVPLSSITTLDTFIALNFTLLEN 653
Db 575 NEIFLTKKMTVCQATSOYYFQSGNEIHVYNDYHHFKTIELDGIATLQTFISLNTSLIEN 634
QY 654 VDFKVIELTRDEKRLSNVFDIETMFREYNYAQRVSGRLKDLDD--LSTNRNQFVDARGS 712
Db 636 IDFASLELYSDEQRASNVFDLEGIFREYNFOAQNIAGRLKDLDAVNSGRNQFVDGLGE 694
QY 713 LMDDLGAVGQTVVNAVSGVATLFSISIVTGFINFKNPFGGMLMIIVIGVLFAYFLTKK 772
Db 696 LMDSLGSGVQSITNLVSTVGGFLFSSLVSGFISFFKNPFGGMLIIVLVAGVILVISLRR 754
QY 773 TKIYETAPIKMIYPEIDKLEREGKSE---IAPISEEBELERIVLAWHIHQNSHMETKTR 829
Db 756 TRMSQQPQVMLYPGIDELAQOAHASGEGPGINPISKTELQAIMLA--LHEQNOEQKRAAQ 812
QY 830 KDPKDSILTTRAQNMLR 847
Db 814 RAAGPSVASRALQAADR 830

```

RESULT 28

```

US-08-475-668-103
; Sequence 103, Application US/08475668
; GENERAL INFORMATION:

```

```
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
; TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING INFLUENZA
; TITLE OF INVENTION: VIRUS TRANSMISSION
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,668
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-026
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEFAX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 103:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 856 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-475-668-103

Query Match 40.48; Score 1839; DB 8; Length 856;
Best Local Similarity 46.18; Pred. No. 1.5e-163;
Matches 368; Conservative 148; Mismatches 240; Indels 42; Gaps 13;

Qy 81 STFPYRVCSAGVGDFVRFQTDHVCPCD-ASDMVHSEGIILYKQNIIPFMRVRYKRVKV 139
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 140 TTSTVYNGIYSDSTNTHQTYFKSIEPWETKMDTIYQCFNSLRLNTGNTLLTYVDRDIN 199
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 200 MTVPFLOPVDGTPDKVKGQPELYLEPGFWGYSRRRTTVNCLMDMFARSPPDFV 259
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 260 TATGDTVEMGPFWSGDEDDHMKHKEPWFVSVINNYKVVDYQNGRTVPLGKTRIFLDREE 319
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 320 YTLISWEHLKN-MSYCPILTUKAFYNGIQTEHGSYHFVANDITASFTTSKEDMKFNNT 378
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 379 YHCLNEIEKAEIEKKYAKYNSTHSHKYCD-LKYFTKDGGLYLVWQPLIONLLDAKN--KL 435
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 436 NNETYG-----RRSRQAESTTDPMMEMTGNAGGEYSSENSI 473
```

```
Db 399 TPTTSPSPSPSPAPSAARGSTPAAVLRERRRDAGNATTP---VPPAPGKSLGTLLNP 455
Qy 474 TVAQVOYAYDNLRIRINNILEDSKAWCREOHRRAALVWNELSKINPTSVMSIMYNPVSA 533
Db 456 ATVOIQFAYDSLRQRINRLMGDLARAWCLEQKQKQNVLRRELTTKINPTTVNSSYIGKAVAA 515
Qy 534 KRIGDVISVNCITVDPQTSVSLHKSURLLSASDEKCFSPVPVTFKFWNDSTIYKGQLGVN 593
Db 516 KRLGDVISVSCQVPVQATVTLKSMR-VPGSETMCSYRPLVSFSEFINDTKTVEGQLGTD 574
Qy 594 NEILLTTLTLETCQENTYFYQAKTDMYIKNYEHLKTVPLSSITLDTFIALNFTLLEN 653
Db 575 NEIFLTKMTEVCQATSYFYQSGNEIHVYNDYHFKTIELDGIATLQTFISLNTSLIEN 634
Qy 654 VDFKVIETLREDEKRLSNVFDIETMPREYNYAQRVSGLRKQLLD-LSTNRNQFVDAFGS 712
Db 635 IDFASLELYSRDEQRASNVFDLEGI FREYNFOAQNIAGLRKOLDNAVSNGRNQFVDGLGE 694
Qy 713 LMDDLGAVGQTVVNAVSGVATLFPSSIVTGFINKPFGGMLMIIVVIGVLFALYFLTKK 772
Db 695 LMDSLGSGVGSITNLVSTVGGLFSSLVSGFISFFKNPFGGMLLVLVAGVVILVISLRR 754
Qy 773 TKIYETAPIKMIYPEIDKLEREGKSE---IAPISSEELERIVLAMHIIHOONSHMETKTR 829
Db 755 TROMSQOPVQMLYPGIDELAAQHASGEGPGCINFIKTELQAIMLA--LHEQNOEQKRAAQ 812
Qy 830 KDPKDSILTRAQNMRLKR 847
Db 813 RAAGPSVASRALQAARDR 830

RESULT 29
US-08-484-223-103
; Sequence 103, Application US/08484223
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
; TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,223
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-029
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEFAX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 103:
; SEQUENCE CHARACTERISTICS:
```

```

; LENGTH: 856 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-484-223-103

Query Match
Best Local Similarity 40.4%; Score 1839; DB 8; Length 856;
Matches 368; Conservative 148; Mismatches 240; Indels 42; Gaps 13;

Qy 81 STFPYRVCSASGVDFRQTHVCPD-ASDMVHSEGLLIYKQNIIPFMRVRKRVKVV 139
Db 44 TSFPRVCELSSHGDLFRFSSDIQCPSTRENHTEGLLMVFKDNIIPYSFKVRSYTKIV 103

Qy 140 TTSTVYNGVSDSIQTHQTFYKSIETPWEKMDTIYOCFNSLRNTGNLLTYVDRDIN 199
Db 104 TNLIYNGVADSVNTHREKFSVDSYETQMDTIYQYNAVKMTKDGLTRVYVDRDGVN 163

Qy 200 MTVPFLQPDVGTDPVKYSGOPELYLEPGFWGSGYRRTTVNCELMDMFARNPPDFV 259
Db 164 ITVNLKPTGGLANGVRRYASOTELYDAPGLIWTYRTTVNCLITDMMAKSNPFDFFV 223

Qy 260 TATGDTVEMSPFSGEDDHENKHEKPFVSVINNYKVVDYQNRGTVPGLKTRIFLDREE 319
Db 224 TTTGQTVEMSPFYDGK--NKETPHERADS FHVRTNYKIYVDYDNRGTNPQERRAFDKGT 281

Qy 320 YTLSEKHLKN-MSYCPLTLWKAIFYNGIOTHEHSGSYHFVANDITASFTTSKEDMKFNTT 378
Db 282 YTLSEKHLKN-MSYCPLTLWKAIFYNGIOTHEHSGSYHFVANDITASFTTSKEDMKFNTT 378

Qy 379 YHCLNEEIKAEIEKKYAKVNSHSHSKYCD-LKYFKTDGGLVLMQPLIQNLLDKN--KL 435
Db 340 FKCIIEQVN-KTHEKYEAVQDRTYTKGOEAITFYITSGGLLAWLPLTPRSLATVKNLTET 398

Qy 436 NNETY-----RRSRQAESTTDPMMMTGNGAGGEYSSENSI 473
Db 399 TPTSPSPSPSPAPSAARGSTPAALVLRRLRRRDAGNATP--VPPTAPGKSLGTLPNP 455

Qy 474 TVAQVQYADNLRIRINILEDLKAWCEQHRALVWNELSKINPTVSMYINRPVSA 533
Db 456 ATVOIQFAYDSLRQINRMGLDARAWCLEQKRONVRLRELTKINPTVMSSIIYKAVAA 515

Qy 534 KRIGDIVISNCIVVDQTSVLSHLKSLRLSLASDEKCFRPPVTFKFMNDSTIYKQGLGVN 593
Db 516 KELGDVISVSCVPVQATVILRKSMR-VPGSETWCYSRPLVSFISFINDTKYEQQLGTD 574

Qy 594 NEILTTTLETQENTYFYFOAKTDMYIYKNYHLKTVPLSSITTLDTFIALNFTLLEN 653
Db 575 NEIFLTKMTEVCQATSOYFQSGNEIHVYNDYHFKTIELDGIATLQTFISLNTSLIEN 634

Qy 654 VDFKVIELYTRDEKLSNVFDIETWFREYNYAQRVSLGRKDLDD-LSTNRNQVDARGS 712
Db 635 IDFALELYSRDEQASNVDFLEGIFREYNFOAQNIAGLRKDLDDNAVNGRNGNQFVDGLGE 694

Qy 713 LMDDLGAQGQTVVNAVSGVATLFSSIVTGFNFINKPFGGLMIIVLIGVLFAIYFLTKK 772
Db 695 LMDSLGSGVQSTNLVSVGGVLFSSLVSGFISFFKNPFGGLMIIVLAVGVILVISLTRR 754

Qy 773 TKIYETAIKMIYPIEDKLKEREKSE---TAPISEELERIVLAMTHIQONSHMETKTR 829
Db 755 TRMSQSQPVQMLYPCIDELAQOAHASGEGGINPIKTELQAIMLA--LHEQNOEQKRAAQ 812

Qy 830 KDPKDSILTRAQNMRLKR 847
Db 813 RAAGSPVASRALQAADR 830

RESULT 30
US-08-484-223A-103
; Sequence 103, Application US/08484223A
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
```

```

; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
; TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 239
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,223A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-029
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 103:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 856 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-484-223A-103

Query Match
Best Local Similarity 40.4%; Score 1839; DB 8; Length 856;
Matches 368; Conservative 148; Mismatches 240; Indels 42; Gaps 13;

Qy 81 STFPYRVCSASGVDFRQTHVCPD-ASDMVHSEGLLIYKQNIIPFMRVRKRVKVV 139
Db 44 TSFPRVCELSSHGDLFRFSSDIQCPSTRENHTEGLLMVFKDNIIPYSFKVRSYTKIV 103

Qy 140 TTSTVYNGVSDSIQTHQTFYKSIETPWEKMDTIYOCFNSLRNTGNLLTYVDRDIN 199
Db 104 TNLIYNGVADSVNTHREKFSVDSYETQMDTIYQYNAVKMTKDGLTRVYVDRDGVN 163

Qy 200 MTVPFLQPDVGTDPVKYSGOPELYLEPGFWGSGYRRTTVNCELMDMFARNPPDFV 259
Db 164 ITVNLKPTGGLANGVRRYASOTELYDAPGLIWTYRTTVNCLITDMMAKSNPFDFFV 223

Qy 260 TATGDTVEMSPFSGEDDHENKHEKPFVSVINNYKVVDYQNRGTVPGLKTRIFLDREE 319
Db 224 TTTGQTVEMSPFYDGK--NKETPHERADS FHVRTNYKIYVDYDNRGTNPQERRAFDKGT 281

Qy 320 YTLSEKHLKN-MSYCPLTLWKAIFYNGIOTHEHSGSYHFVANDITASFTTSKEDMKFNTT 378
Db 282 YTLSEKHLKN-MSYCPLTLWKAIFYNGIOTHEHSGSYHFVANDITASFTTSKEDMKFNTT 378

Qy 379 YHCLNEEIKAEIEKKYAKVNSHSHSKYCD-LKYFKTDGGLVLMQPLIQNLLDKN--KL 435
Db 340 FKCIIEQVN-KTHEKYEAVQDRTYTKGOEAITFYITSGGLLAWLPLTPRSLATVKNLTET 398

Qy 436 NNETY-----RRSRQAESTTDPMMMTGNGAGGEYSSENSI 473
Db 399 TPTSPSPSPSPAPSAARGSTPAALVLRRLRRRDAGNATP--VPPTAPGKSLGTLPNP 455

Qy 474 TVAQVQYADNLRIRINILEDLKAWCEQHRALVWNELSKINPTVSMYINRPVSA 533
Db 456 ATVOIQFAYDSLRQINRMGLDARAWCLEQKRONVRLRELTKINPTVMSSIIYKAVAA 515

Qy 534 KRIGDIVISNCIVVDQTSVLSHLKSLRLSLASDEKCFRPPVTFKFMNDSTIYKQGLGVN 593
Db 516 KELGDVISVSCVPVQATVILRKSMR-VPGSETWCYSRPLVSFISFINDTKYEQQLGTD 574

Qy 594 NEILTTTLETQENTYFYFOAKTDMYIYKNYHLKTVPLSSITTLDTFIALNFTLLEN 653
Db 575 NEIFLTKMTEVCQATSOYFQSGNEIHVYNDYHFKTIELDGIATLQTFISLNTSLIEN 634

Qy 654 VDFKVIELYTRDEKLSNVFDIETWFREYNYAQRVSLGRKDLDD-LSTNRNQVDARGS 712
Db 635 IDFALELYSRDEQASNVDFLEGIFREYNFOAQNIAGLRKDLDDNAVNGRNGNQFVDGLGE 694

Qy 713 LMDDLGAQGQTVVNAVSGVATLFSSIVTGFNFINKPFGGLMIIVLIGVLFAIYFLTKK 772
Db 695 LMDSLGSGVQSTNLVSVGGVLFSSLVSGFISFFKNPFGGLMIIVLAVGVILVISLTRR 754

Qy 773 TKIYETAIKMIYPIEDKLKEREKSE---TAPISEELERIVLAMTHIQONSHMETKTR 829
Db 755 TRMSQSQPVQMLYPCIDELAQOAHASGEGGINPIKTELQAIMLA--LHEQNOEQKRAAQ 812

Qy 830 KDPKDSILTRAQNMRLKR 847
Db 813 RAAGSPVASRALQAADR 830
```

```
Db 399 TPTSSPPSPAPSAARGSTPAALVLRRRRRDAGNATTP---VPPTAPGKSLGTLNPP 455
Qy 474 TVAQOYAYDNLRIIRINNILEDLSKAWCREOHRAALVWNLKINPTSVMSIYNRPVSA 533
Db 456 ATVOIQFAYDSLRQINRMGLDGLARAWCLEQKQNMVLRKLTINPTTVWSSIYGKAVAA 515
Qy 534 KRIGDVISVNCIIVVDQTSVSLHSLRLSASDEKCFSPRPVTFKFWNDSTIYKQGLGVN 593
Db 516 KRLGDVISVOCVPVNOATVTLRKSMT-VPGETWCYSRPLVSFINDTKTYEGQGLTD 574
Qy 594 NEILLTTTILETCQENTYFYFOAKTDMYIYKNYBHLKTVPLSSITTLDTFIALNFTLLEN 653
Db 575 NEIFLTKMTEVCQATSYFYFQSGNEIHVYNDYHFKTIELDGIATLQTFISLNTSLIEN 634
Qy 654 VDFKVIELYTRDEKRLSNVFDIETMFREYNYAQRVSLGRKDLDD--LSTNRNQFVDAFGS 712
Db 635 IDFASLELYSRDEQRASNVFDLEGIFREYNFQAQNIAGRLKDLNNAVSNRNOFVDGLGE 694
Qy 713 LMDDLGAVGQTVVNAVSGVATLFSISVTGFINFKPFGGMLMIIVVIGVLFAYIELTKK 772
Db 695 LMDSLGVSQGITNLVSTVGGFLSLSVSGFISFFKNPFGGMLLVLVAGVILVISLRR 754
Qy 773 TKIYETAPIKMIYPEIDKLKEREKSE---IAPISEEELERIVLAMIHQNSHMETKTR 829
Db 755 TRQMSQQPVQMLYPCIDELAQQHASGEGPGINPISKTELQAIMLA--LHEQNOEQKRAAQ 812
Qy 830 KDPKDSILTRAQNMURKR 847
Db 813 RAAGPSVASRALQAARDR 830

RESULT 31
US-08-484-223D-103
; Sequence 103, Application US/08484223D
; GENERAL INFORMATION:
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; TITLE OF INVENTION: ISOLATED PEPTIDES DERIVED FROM HUMAN
; TITLE OF INVENTION: IMMUNODEFICIENCY VIRUS TYPES 1 AND 2 CONTAINING
; FILE REFERENCE: 7872-029-999
; CURRENT APPLICATION NUMBER: US/08/484,223D
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 08/470,896
; PRIOR FILING DATE: 1995-06-06
; PRIOR APPLICATION NUMBER: 08/360,107
; PRIOR FILING DATE: 1994-12-20
; PRIOR APPLICATION NUMBER: 08/255,208
; PRIOR FILING DATE: 1994-06-07
; PRIOR APPLICATION NUMBER: 08/073,028
; PRIOR FILING DATE: 1993-06-07
; NUMBER OF SEQ ID NOS: 272
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 103
; LENGTH: 856
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
US-08-484-223D-103

Query Match 40.4%; Score 1839; DB 8; Length 856;
Best Local Similarity 46.18; Pred. No. 1.5e-163;
Matches 368; Conservative 148; Mismatches 240; Indels 42; Gaps 13;

Qy 81 STFPYRVCSAGVGDFRFDTHVCPD-ASDMVHSEGILLIYKQNIIPFMRVRYKRVV 139
Db 44 TSFPYRVCELSSHGDLFRFSSDIQCPGFTRENHTEGLLMVFKDNIIPYSFKVRSYTKIV 103
Qy 140 TTSTVNGIYSDSITNOHTYKSIPEWETEMDTIYQCFNSRLNCGNLLTVYDRDDIN 199
Db 104 TNLIVNGWADSVTRNHEEKFSVDSYETQDMPTIYQCYNAVWTKDKGLTRVYVDRDGVN 163
Qy 200 MTVFLQPDVGVTPDVKRYGSGQPELYLEPGFWGWSYRRRTTNCVCLMDMFARSPPDFV 259
```

```
Db 164 ITVNLKPTGGLANGVRYASQIELYDAPGWLWITYRTRITVNCILITDMAKSNSPDFV 223
Qy 260 TATGDTVEMSPFWSGDEDDHENKMKHPFVSVVNNYKVDYQNRGTVPGLKTRIFLDREE 319
Db 224 TTTGTQVEMSPFFDVGK--NKETFERADSPHVRTNYKIVDYDNRGTPNQGERAFLDKGT 281
Qy 320 YTLISWEKHLKN-MSYCPILTLKAFYNGIQTEHSGSVHYFVANDITASFTTSKEDMKERN 378
Db 282 YTLISWK--LENRTAYCPLOHWOFTDSTIATETGKSIHFVTDGTSFVNTTGVIELPDA 339
Qy 379 YHCLNEIEIKAEIKYAKVNSTHSHKYGD-LKYFKTDGGLYLVWQPLIQNLRLDKN--KL 435
Db 340 FKICIEOVN-KTHEKYEAVQDRTYKQEBALTYFITSGLLLAWLPLTPRSLATVKNLTTEL 398
Qy 436 NNETY-----RRSRQAESTTDDPMWMTGNGAGGYSSENSI 473
Db 399 TPTSSPPSPAPSAARGSTPAALVLRRRRRDAGNATTP---VPPTAPGKSLGTLNPP 455
Qy 474 TVAQOYAYDNLRIIRINNILEDLSKAWCREOHRAALVWNLKINPTSVMSIYNRPVSA 533
Db 456 ATVOIQFAYDSLRQINRMGLDGLARAWCLEQKQNMVLRKLTINPTTVWSSIYGKAVAA 515
Qy 534 KRIGDVISVNCIIVVDQTSVSLHSLRLSASDEKCFSPRPVTFKFWNDSTIYKQGLGVN 593
Db 516 KRLGDVISVOCVPVNOATVTLRKSMT-VPGETWCYSRPLVSFINDTKTYEGQGLTD 574
Qy 594 NEILLTTTILETCQENTYFYFOAKTDMYIYKNYBHLKTVPLSSITTLDTFIALNFTLLEN 653
Db 575 NEIFLTKMTEVCQATSYFYFQSGNEIHVYNDYHFKTIELDGIATLQTFISLNTSLIEN 634
Qy 654 VDFKVIELYTRDEKRLSNVFDIETMFREYNYAQRVSLGRKDLDD--LSTNRNQFVDAFGS 712
Db 635 IDFASLELYSRDEQRASNVFDLEGIFREYNFQAQNIAGRLKDLNNAVSNRNOFVDGLGE 694
Qy 713 LMDDLGAVGQTVVNAVSGVATLFSISVTGFINFKPFGGMLMIIVVIGVLFAYIELTKK 772
Db 695 LMDSLGVSQGITNLVSTVGGFLSLSVSGFISFFKNPFGGMLLVLVAGVILVISLRR 754
Qy 773 TKIYETAPIKMIYPEIDKLKEREKSE---IAPISEEELERIVLAMIHQNSHMETKTR 829
Db 755 TRQMSQQPVQMLYPCIDELAQQHASGEGPGINPISKTELQAIMLA--LHEQNOEQKRAAQ 812
Qy 830 KDPKDSILTRAQNMURKR 847
Db 813 RAAGPSVASRALQAARDR 830

RESULT 32
US-08-485-551-103
; Sequence 103, Application US/08485551
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING INFLUENZA VIRUS
; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```

Db      695  LMDSLGSGVQSIINLVSTVGGLESSLVSGFISPFKNPFGGMLILVLVAGVVILVLSLRR 755
Qy      773  TKIYETAPIKWIYPEIDKLKEREKSE---IAPISEBELEERIVLAMHHQONSHMETKTR 829
Db      755  TROMSQOPVQMLYFGIDELAQOAHSGGPGINPISKTELQAIMLA--LHEQNEQEKRAAQ 812
Qy      830  KDPKDSILTILTRAQNWLRKR 847
Db      813  RAAGPSVASRALQAARDR 830

RESULT 33
US-08-487-266-103
; Sequence 103, Application US/08487266
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING HIV TRANSMISSION
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,266
; FILING DATE: 07-JUN-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-025
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 103:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 856 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-487-266-103

Query Match 40.4%; Score 1839; DB 8; Length 856;
Best Local Similarity 46.1%; Pred. No. 1.5e-163;
Matches 368; Conservative 148; Mismatches 240; Indels 42; Gaps 13;

Qy      81  STFPYRVCASGVDYFRFQTDHVCDD-ASDMVHSSGILLIYKONIIPMFVRVKYRKVV 139
Db      44  TGFPPFRCVLSHGDUFRFSSDIQCPSFGRENHTEGLMLVFDKNIILPYSPKVSRYTKIV 103
Qy      140  TTSTVYNGIVSDITNQHTPFYKSIEPWETKMDTIYQCENSLRINTCGNLLTYVDRDDIN 199
Db      104  TWILIYNGWYADSVTNRHEKFSVDSYETDMDTIYQCYNVAVKMTKDLTRVYVDRDGVN 163
Qy      200  MTVFLQDPDQVTPDVVKRYGSGQPELYLPEPGFWGWSGRRRTTVNCVELMDMWFARSNPPFFV 259
Db      164  ITVNLKPTGGLANGVRYASOTELYDAPGWLITWYTRTRTVNCLITDMMAKSNPSPFFV 223

```

```
QY 260 TATGDTVEMSPFWSGEDDHENKMKHKEPWFVSVINNYKVVDYQNRGTVPGLCKTRIFLDREE 319
DB 224 TTTGQTVEMSPFVDGK--NKETHRADSHVRTNYKIVDYDNRGNTNPOGERRAFDLKGT 281
QY 320 YTLSEWKLKN-MSYCPLTLWKAFYNGIQTEHSGSYHFVANDITASFTTSKEDMKEFNTT 378
DB 282 YTLWK--LENRTAYCPLQHWQTFDSTIATETGKSIHFVDEGTSFVNTTVGIELPDA 339
QY 379 YHCLNEIKAEIKKAKYNSTHSHKYGD-LKYFKTDGGLYLVQPLIQNRLLDAKN--KL 435
DB 340 FKCI EQVN-KTHEKYEAVQDRYTKQEAITYFITSGLLLAWLPLTPRSLATVKNLTEL 398
QY 436 NNETYS-----RSRROAESTTDPMMEMTGNAGGEYSSENSI 473
DB 399 TPTSPSPSPSPAPSAARGSTPAVLRRRRRDAGNATTP---VPPTAPGKSLGTINNP 455
QY 474 TVAQVQAYDNLRIIRINILEDLSKAWCREQHRAALVWNLKSKINPTSVMSMITYNRPVSA 533
DB 456 ATVQIQFAYDSLRRQINRMGLDLARAWCLEQKQNMVLRRLTKINPTVMSSYIGKAVAA 515
QY 534 KRIGDVISVNCIVVDQTSVLSLHKSRLLSASDEKCFSPRPVTFKFMNDSTIYKQGLGVN 593
DB 516 KRLGDVISVQCVPVNOATVTLRKSMR-VPGSETMCSYRPLVSFSFINDTKTYEGOLGTD 574
QY 594 NEILLTTTLETQENTYFYFOAKTDMYIYKNYEHKTVPLSSITTLDTFIALNFTLLEN 653
DB 575 NEIFLTKKMTEVCOATSYQYFQSGNEIHVYNDYHFKTIELDGIATLQTFISLNTSLIEN 634
QY 654 VDFKVIETRLDEKRLSNVFDIETMFREYNYAQRVSLGRKOLL--LSTNRNQFVDAFGS 712
DB 635 IDPASLELYSRDEQRASNVFDLEGI FREYNFQAQNIAGRLKDLNNAVSNRGNQFVDGLGE 694
QY 713 LMDDLGAVGQTVVNAVSGVATLFSSIVTGFINKPFGGMLMIIVVIGVLFALYELTKK 772
DB 695 LMDSLGSGVGSQITNLVSTVGGLFSSLVSGFISFPKPNFGGMLILVLVAGVILVLSLRR 754
QY 773 TKIYETAPIKMIYPEIDKLKEREKSE--IAPISSEELERIVLAMHIHQONSHMETKTR 829
DB 755 TROMSQOPVOMLYPGIDELAQHRSAGGPGINPLSKTELQAIMLA--LHEQNOEQKRAAQ 812
QY 830 KDPKDSILTRAQNMRLRX 847
DB 813 RAAGPSVASRALQAARD 830

RESULT 34
US-08-487-266A-103
; Sequence 103, Application US/08487266A
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING HIV TRANSMISSION
; NUMBER OF SEQUENCES: 239
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
```

```
; APPLICATION NUMBER: US/08/487,266A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-025
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 103:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 856 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-487-266A-103

Query Match 40.4%; Score 1839; DB 8; Length 856;
Best Local Similarity 46.1%; Pred. No. 1.5e-163;
Matches 368; Conservative 148; Mismatches 240; Indels 42; Gaps 13;

QY 81 STPPYRVCASGVGDVRFQTDHVCDD-ASDMVHSEGIILLIYKQNIIPFMFRVRYKRVK 139
DB 44 TSFPFVVCVLSHGDLFPRESSDIQCFSGFTRENHTEGLLMVFKDNIIPYSFKVRSYTKIV 103
QY 140 TTSVTVNGIYSISINQHTFYKSIPEWTEKMDTIYOCFNSRLNTGGNLLTYVDRDDIN 199
DB 104 TNILYNGWYADSVTNRHEEKFSVDSYETQDMDTIYQCYNVAKMTKDGLTRVYVDRDGVN 163
QY 200 MTVFLOPVDGVTDPVKYCSQPELYLERPGFWGVSRRRTTVNCELMDMFARSNPPDFV 259
DB 164 ITNLKPTGCLANGVARYASQTELYDAPGLWLTWYTRTIVNCLITDMMAKSNPFDFFV 223
QY 260 TATGDTVEMSPFWSGEDDHENKMKHKEPWFVSVINNYKVVDYQNRGTVPGLCKTRIFLDREE 319
DB 224 TTTGQTVEMSPFVDGK--NKETHRADSHVRTNYKIVDYDNRGNTNPOGERRAFDLKGT 281
QY 320 YTLSEWKLKN-MSYCPLTLWKAFYNGIQTEHSGSYHFVANDITASFTTSKEDMKEFNTT 378
DB 282 YTLWK--LENRTAYCPLQHWQTFDSTIATETGKSIHFVDEGTSFVNTTVGIELPDA 339
QY 379 YHCLNEIKAEIKKAKYNSTHSHKYGD-LKYFKTDGGLYLVQPLIQNRLLDAKN--KL 435
DB 340 FKCI EQVN-KTHEKYEAVQDRYTKQEAITYFITSGLLLAWLPLTPRSLATVKNLTEL 398
QY 436 NNETYS-----RSRROAESTTDPMMEMTGNAGGEYSSENSI 473
DB 399 TPTSPSPSPSPAPSAARGSTPAVLRRRRRDAGNATTP---VPPTAPGKSLGTINNP 455
QY 474 TVAQVQAYDNLRIIRINILEDLSKAWCREQHRAALVWNLKSKINPTSVMSMITYNRPVSA 533
DB 456 ATVQIQFAYDSLRRQINRMGLDLARAWCLEQKQNMVLRRLTKINPTVMSSYIGKAVAA 515
QY 534 KRIGDVISVNCIVVDQTSVLSLHKSRLLSASDEKCFSPRPVTFKFMNDSTIYKQGLGVN 593
DB 516 KRLGDVISVQCVPVNOATVTLRKSMR-VPGSETMCSYRPLVSFSFINDTKTYEGOLGTD 574
QY 594 NEILLTTTLETQENTYFYFOAKTDMYIYKNYEHKTVPLSSITTLDTFIALNFTLLEN 653
DB 575 NEIFLTKKMTEVCOATSYQYFQSGNEIHVYNDYHFKTIELDGIATLQTFISLNTSLIEN 634
QY 654 VDFKVIETRLDEKRLSNVFDIETMFREYNYAQRVSLGRKOLL--LSTNRNQFVDAFGS 712
DB 635 IDPASLELYSRDEQRASNVFDLEGI FREYNFQAQNIAGRLKDLNNAVSNRGNQFVDGLGE 694
QY 713 LMDDLGAVGQTVVNAVSGVATLFSSIVTGFINKPFGGMLMIIVVIGVLFALYELTKK 772
DB 695 LMDSLGSGVGSQITNLVSTVGGLFSSLVSGFISFPKPNFGGMLILVLVAGVILVLSLRR 754
QY 773 TKIYETAPIKMIYPEIDKLKEREKSE--IAPISSEELERIVLAMHIHQONSHMETKTR 829
```



```
Db 755 TRMSQPPVQMLYPGIDELAQHASGEGPGINPISKTELQAIMLA--LHEQNEQKRAAQ 812
Qy 830 KDPKDSILTRAQNMRLKR 847
Db 813 RAAGPSVASRALQAADR 830

RESULT 35
US-08-487-355-103
; Sequence 103, Application US/08487355
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING HEPATITIS B VIRUS
; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,355
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-027
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 103:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 856 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-487-355-103

Query Match 40.4%; Score 1839; DB 8; Length 856;
Best Local Similarity 46.1%; Pred. No. 1.5e-163;
Matches 368; Conservative 148; Mismatches 240; Indels 42; Gaps 13;

Qy 81 STFPYRVCSAGVGDFVQTDHVCPCD-ASDMVHSEGIILLYKONIIPMFVRVRYKRV 139
Db 44 TSFPRVVELSHGDLFRSSDIQPCPSGTRENHTGGLLMVFKONIIPYFKVRSYTKIV 103
Qy 140 TTSTVYNGYSDITNQHTFFKYSIEPWETERKMDTIYQCNSLRNLNTGGNLLTYVDRDDIN 199
Db 104 TWILIYNGWYADSVNTRHEEKFSVDSYETDQMDTIYQCVNAVMTKOGITRYVYVDRDGVN 163
Qy 200 MTFVLPQVDGVTVPDKRVQSGDELYLEPCGFWGSRRTTNCVCELMDFARSNPPDFPV 259
Db 164 ITVNLKPTGGLANGVRRYASOTELYDAPGWLITVTRTNTCLITDMMAKSNSPFDVF 223
Qy 260 TATGDTVEMSPWMSGDDHKNKHEKPFWFSVNNYKVVQYQNRCTVPLGKTRIFLDREE 319

Db 224 TTTGQTVEMSPFYDGK--NKETPHERADSPHVRNYKIVDYDNRGTNPQGERRAFLDKGT 281
Qy 320 YTLSEWEXHLKN-MSYCPCLTLWKAFYNGIQTEHSSSYHYFVANDITASFTTSKEDKKEFWNTT 378
Db 282 YTLSSWK--LENRTAYCPLOHMQTFDSTIATETGSKIHFVTDDEGTSSFTVNTTVGIELPDA 339
Qy 379 YHCLNEEIKAEIKKAKVNSTHSHKYGD-LKYFXTDGGLYLVQPLQNLQNRLDARN--KL 435
Db 340 FKCIEBQVN-KTHEKEYAVQDRYTKQEAITYFITSGLLLAWLPLPRSLATVKNLTEL 398
Qy 436 NNETY-----RRSRQAESTTDPMMEMTNGACGEYSSENSI 473
Db 399 TPTTSSPPSSPPSAARGSTPAAVLRRRRRDAGNATP---VPPTAPGKSLGTLNPN 455
Qy 474 TVAQVOYAYDNLRIRINNILEDLSKANCREQHRALVWNELSKINPTSVMSIYNRPVSA 533
Db 456 ATVOIQFAYDSLRQINRMGLDLARAWCLEQKQNMVLRRLTKINPTTVMSSIYGVKAVAA 515
Qy 534 KRIGDVISVNCIVVDOTSVSLHKSRLLSASDEKCFSRPPVTFKFMNDSTIYKQGLGVN 593
Db 516 KRLGDVISVQCVPNQATVTLRKSMT--VPGSETWCYSRPLVSFSFINDTKTYEGQLGTD 574
Qy 594 NEILLTTTLETQCENTEYYPQAKTDMYIYKNYEHKLTVPPLSSITTLDTFIALNFTLLEN 653
Db 575 NEIFLTKKTEVCQATSQYFYQSGNEIHVYNDYHHFKTIELDGIATLQTFISLNTSLIEN 634
Qy 654 VDFKVIELYTRDEXRLSNVFDIETMFREYNYAQRVSGRLRKDLDD--LSTNRNQFVDARGS 712
Db 635 IDFASLELYSDEQRASNVFDELGIFREYNFQAQIAGLRKDLDNVNSNGRNQFVDGLGE 694
Qy 713 LMDDLGAVGQTVNNAVSVATLRFSSIVTGRFNETKNPFGGMGLIIVIGLVFAIYFLTKK 772
Db 695 LMDSLGVGQSITNLVSTVGGFLSVSGFISFPNPFPGMGLIIVLVAGVVILVISLRR 754
Qy 773 TKIYETAPIKMIYPEIDKLKEREKSE---IAPISEEELERIVLAMHIHQONSHMETKTR 829
Db 755 TRMSQPPVQMLYPGIDELAQHASGEGPGINPISKTELQAIMLA--LHEQNEQKRAAQ 812
Qy 830 KDPKDSILTRAQNMRLKR 847
Db 813 RAAGPSVASRALQAADR 830

RESULT 36
US-08-487-355A-103
; Sequence 103, Application US/08487355A
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING HEPATITIS
; TITLE OF INVENTION: B VIRUS TRANSMISSION
; NUMBER OF SEQUENCES: 273
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,355A
```

```

Db      755 TRQMSQPQVQLYICGIDELAQOAHSGEGPGINPISKTELQAIMLA--LHEQNQEQKRAAQ 81
Qy      830 KDPKXDSILTRAQNMLRKR 847
      :      :      :      :      :
Db      813 RAAGPSVASRALQAARDR 830

RESULT 37
US-08-919-600-103
: Sequence 103, Application US/08919600
: GENERAL INFORMATION:
: APPLICANT: Bolognesi, Dani P.
: APPLICANT: Matthews, Thomas J.
: APPLICANT: Wild, Carl T.
: APPLICANT: Barney, Shawn O.
: APPLICANT: Lambert, Dennis M.
: APPLICANT: Petteway, Stephen R.
: APPLICANT: Langlois, Alphonse J.
: TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION
: TITLE OF INVENTION: OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
: TITLE OF INVENTION: TRANSMISSION
: NUMBER OF SEQUENCES: 273
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Pennie & Edmonds
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: USA
: ZIP: 10036-2711
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/919,600
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/470,896
: FILING DATE: 06-JUN-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Coruzzi, Laura A.
: REGISTRATION NUMBER: 30,742
: REFERENCE/DOCKET NUMBER: 7872-020
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 790-9090
: TELEFAX: (212) 869-9741/8864
: TELEX: 66141 PENNIE
: INFORMATION FOR SEQ ID NO: 103:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 856 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: unknown
: MOLECULE TYPE: protein
: US-08-919-600-103

Query Match 40.4%; Score 1839; DB 13; Length 856;
Best Local Similarity 46.1%; Pred. No. 1.5e-163;
Matches 368; Conservative 148; Mismatches 240; Indels 42; Gaps 13

Qy      81 STPYRVCASGVGDVRFQTDHVCPD-ASDMVHSEGLILLYKQNIIFPMFVRVKYKV 139
      :      :      :      :      :      :      :      :      :      :
Db      44 TSFPRVCYELSSHGDLRFSSDIQCP8FGTRENTEGLLMVFKONIIPYSFKVRSYTKIV 103
      :      :      :      :      :      :      :      :      :      :
Qy      140 TTSTVYNGIYSDSITNQHTFYKSTPEWETEKMDIYQCFNSLRNLNTGNNLLTYVDRDDIN 199
      :      :      :      :      :      :      :      :      :      :
Db      104 TNLINGWADSVTNRHEEKFSVDSVETQDMDIYQCYNNAVKMTKDGILTRVYVDRDGVN 163
      :      :      :      :      :      :      :      :      :      :
Qy      200 MTFVLQPDVGDTPDKRYGSGQPELYLEFGFWGVSRYRRTTVCNCELMDFEARNPPFPFV 259
      :      :      :      :      :      :      :      :      :      :
Db      164 ITNLKPTGGLANGVRVYASOTELYDAPGLIWTYRTRTTVNCILITDMMAKSNSPFDFFV 223
      :      :      :      :      :      :      :      :      :      :

```

```
QY 260 TATGDTVEMSPFSGEDDHNKMKHEKPFVSVNNYKVVDYQNRGTVPGLKTRIFLDREE 319
Db 224 TTTGQTVEMSPFYDGGK--NKETPHERADSFHVRTNYKIVDYNRGTPNQGERRAFLDKGT 281
QY 320 YTLSEKHLKN-MSYCPDLTWKAFYNGIOTHSYGHFVANDITASFSTTSKEDMKFNTT 378
Db 282 YTLSEKHLKN-MSYCPDLTWKAFYNGIOTHSYGHFVANDITASFSTTSKEDMKFNTT 378
QY 379 YHCLNEEIKAEIEKKYAKVNSTHSGYGD-LKYFKTDGGLVLMQPLIQNRLLDKN--KL 435
Db 340 FKCIIEQVN-KTHEKEAVQDRYTKGQEAITYFITSGLLLAWLPLTPRSLATVKNLTJEL 398
QY 436 NNETYSS-----RRSRQAESTTDPMMEMTNGAGGEYSSENSI 473
Db 399 TTTPTSSPPSPSPAPSAARGSTPAVLRRRRRDAGNATP---VPPTAPGKSLGTNNP 455
QY 474 TVAQOVAYDNLRIIRINILEDLKAWCREQHRAALVWNLKINPTSVMSIYNRPVSA 533
Db 456 ATVOIQFAYDSLRRQINRMGLDLARAWCLEQRQNMVRLTKINPTVMSSYKAVAA 515
QY 534 KRIGDIVSNCIIVDQTSVSLHKLRLLSASDEKCFRPPVTEKFMNDSTIYKGLGVN 593
Db 516 KRLGDIVSQCVPVNOATVTLRKSMR-VPGSETWCYSRPLVSPFINDTKYEGQGLTD 574
QY 594 NEILLTTTLETQENTYFYFOAKTDMYIKYKVEHLKTVPLSSITTLDTFIALNFTLLEN 653
Db 575 NEIFLTKKMTVECOATSOYFYFQSGNEIHVYNDYHHFKTIELDGIATLQTFISLNTSJEN 634
QY 654 VDFKVIELYTRDEKRLSNVFDIETMPREYNYAQRVSGRLKDLDD-LSTNRNQVDAFGS 712
Db 635 IDFASLELYSRDEQRASNVFDLEGIIFREYNFQAQNIAGLRKDLDNAVNSGRNQVDGLGE 694
QY 713 LMDDLGAVGQTVNNAVSGVATLFSSIVTGFNFINKNPFPGMLIIVIGVLFAIYFLTKK 772
Db 695 LMDSLGSGVQSTINLVSTVGGFLSSVSGFISFFKNPFPGMLILVLVAGVILVISLRR 754
QY 773 TKIYETAPIKMIYPIBKLEKREGKSE---IAPISEELERIVLAWHHQOONSHMETKTR 829
Db 755 TQMSQOPQOMLYPGIDELAQHASGEFGINPISKTELQAIMLA--LHEQNOEQKRAQ 812
QY 830 KDPKDSILTRAQNMRLKR 847
Db 813 RAAGPSVASRALQAARDR 830
```

RESULT 38

US-09-502-445-103

; Sequence 103, Application US/09502445

; GENERAL INFORMATION:

; APPLICANT: Bolognesi, Dani P.

; APPLICANT: Matthews, Thomas J.

; APPLICANT: Wild, Carl T.

; APPLICANT: Barney, Shawn O.

; APPLICANT: Lambert, Dennis M.

; APPLICANT: Langlois, Alphonse J.

; TITLE OF INVENTION: METHODS FOR INHIBITION OF

; TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV

; NUMBER OF SEQUENCES: 244

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pennie & Edmonds LLP

; STREET: 1155 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10036-2711

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/502,445

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/487,266

; FILING DATE: 07-JUN-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Coruzzi, Laura A.

; REGISTRATION NUMBER: 30,742

; REFERENCE/DOCKET NUMBER: 7872-025

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 790-9090

; TELEFAX: (212) 869-9741/8864

; TELEX: 66141 PENNIE

; INFORMATION FOR SEQ ID NO: 103:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 856 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: unknown

; MOLECULE TYPE: protein

US-09-502-445-103

Query Match 40.4%; Score 1839; DB 19; Length 856;

Best Local Similarity 46.1%; Pred. No. 1.5e-163;

Matches 368; Conservative 148; Mismatches 240; Indels 42; Gaps 13;

QY 81 STFPYRVCSAGVGDFRFDTHVCPD-ASDMVHSEGLLIYKQNIIPFMFRVRYKRVV 139

Db 44 TSFPFRVCELSSHGDLFRFSSDIQCPSPGTRENHTEGLLMVFKDNIIPYSFKVRSYTKIV 103

QY 140 TTSIVYNGIYSDSTNQHTFYKSIETPWEKMDIYQCFNSRLRLNTGNTLLTYVDRDDIN 199

Db 104 TNLIIYNGWYADSVNRHEEKFSVDYETDQMDTIYQCYNAVAKMTKDLTRVYVDRDGVN 163

QY 200 MTVLQPLQVDGVTDPVKRYGQPELYLPBGWPMGVSRRRTVNCCLMDMFARSNPDPDFV 259

Db 164 ITVNLKPTGGLANGVRRYASQTELYDAPGLWITVTRRTVNCCLITDMAKSNPDPDFV 223

QY 260 TATGDTVEMSPFSGEDDHNKMKHEKPFVSVNNYKVVDYQNRGTVPGLKTRIFLDREE 319

Db 224 TTTGQTVEMSPFYDGGK--NKETPHERADSFHVRTNYKIVDYNRGTPNQGERRAFLDKGT 281

QY 320 YTLSEKHLKN-MSYCPDLTWKAFYNGIOTHSYGHFVANDITASFSTTSKEDMKFNTT 378

Db 282 YTLSEKHLKN-MSYCPDLTWKAFYNGIOTHSYGHFVANDITASFSTTSKEDMKFNTT 378

QY 379 YHCLNEEIKAEIEKKYAKVNSTHSGYGD-LKYFKTDGGLVLMQPLIQNRLLDKN--KL 435

Db 340 FKCIIEQVN-KTHEKEAVQDRYTKGQEAITYFITSGLLLAWLPLTPRSLATVKNLTJEL 398

QY 436 NNETYSS-----RRSRQAESTTDPMMEMTNGAGGEYSSENSI 473

Db 399 TTTPTSSPPSPSPAPSAARGSTPAVLRRRRRDAGNATP---VPPTAPGKSLGTNNP 455

QY 474 TVAQOVAYDNLRIIRINILEDLKAWCREQHRAALVWNLKINPTSVMSIYNRPVSA 533

Db 456 ATVOIQFAYDSLRRQINRMGLDLARAWCLEQRQNMVRLTKINPTVMSSYKAVAA 515

QY 534 KRIGDIVSNCIIVDQTSVSLHKLRLLSASDEKCFRPPVTEKFMNDSTIYKGLGVN 593

Db 516 KRLGDIVSQCVPVNOATVTLRKSMR-VPGSETWCYSRPLVSPFINDTKYEGQGLTD 574

QY 594 NEILLTTTLETQENTYFYFOAKTDMYIKYKVEHLKTVPLSSITTLDTFIALNFTLLEN 653

Db 575 NEIFLTKKMTVECOATSOYFYFQSGNEIHVYNDYHHFKTIELDGIATLQTFISLNTSJEN 634

QY 654 VDFKVIELYTRDEKRLSNVFDIETMPREYNYAQRVSGRLKDLDD-LSTNRNQVDAFGS 712

Db 635 IDFASLELYSRDEQRASNVFDLEGIIFREYNFQAQNIAGLRKDLDNAVNSGRNQVDGLGE 694

QY 713 LMDDLGAVGQTVNNAVSGVATLFSSIVTGFNFINKNPFPGMLIIVIGVLFAIYFLTKK 772

Db 695 LMDSLGSGQITNLVSTVGLFSSLVSGFISFFKNPFGMLILVLVAGVILVSLTR 754
Qy 773 TKIVETAPIKMIYPEIDKLKEREKSE---IAPISEEELERIVLAMIHOONSHMETKTR 829
Db 755 TROMSQOPVQMLYPGIDELAQQHASGEGPGINPISKTELQAIMLA--LHEQNOEQKRAAQ 812
Qy 830 KDPKDSILTRAQNMRLKR 847
Db 813 RAAGPSVASRALQAARDR 830

RESULT 39
US-10-267-682-103
; Sequence 103, Application US/10267682
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; Matthews, Thomas J.
; Wild, Carl T.
; Barney, Shawn O.
; Lambert, Dennis M.
; Petteway, Stephen R.
; Langlois, Alphonse J.

; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
; MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
; TRANSMISSION
; NUMBER OF SEQUENCES: 239
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/267,682
; FILING DATE: 08-Oct-2002
; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,223A
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-029

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE

; INFORMATION FOR SEQ ID NO: 103:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 856 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: unknown

; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 103:
US-10-267-682-103

Query Match 40.4%; Score 1839; DB 28; Length 856;
Best Local Similarity 46.1%; Pred. No. 1.5e-163;
Matches 368; Conservative 148; Mismatches 240; Indels 42; Gaps 13;

Qy 81 STPPYRVCASGVGDVFRFQTDHVCPPD-ASDMVHSEGLIILYKONIIIPFMRVKRKVV 139
Db 44 TSPFPRVCSHGGDLFRFSDIQCPFGTRENHTEGLLMVFKDNIIPYGFVKRSYTKIV 103
Qy 340 TTSTVYNGIVSDSITNQHTFYKSTPEWETEKMDTIYQCFNSRLNTGGLNTYVDRDDIN 199

Db 104 TNLINYGWADSVTRHEEKFSDVSYETQMDTIYQCYNAVMTKDGRLTRVYVDRGVN 163
Qy 200 MTVFLQPDVGTDPVKRYGSOPELYLEPGWFWGYSRRRTTVNCCLMDPMFARSNPPDFV 259
Db 164 ITVNLKPTGGLANGVRRYASOTELYDAPGWLITWYTRTRTTVNCCLITDMAKSNPPDFV 223
Qy 260 TATGDTIVEMSPFWSGEDDHNKHEKFWFVSVNNYKVVDYQNRGTVPLGKTRIFLDREE 319
Db 224 TTTGQTVEMSPFYDGK--NKETFHERADSPHVRTNYKIVDYDNRGTNPQGERRAFLDKGT 281
Qy 320 YTLISWEKHLKN-MSYCPCLTLWKAFYNGIQTEHSGSYHFVANDITASTFTTSKEDMKEFNTT 378
Db 282 YTLISWK--LENRTAYCPLQHWQTFDSTIATETGKSHFVTDEGTSSTVTTTVCIELPDA 339
Qy 379 YHCLNEIKAEIEKKYAKVNSTHSHKYGD-LKYFKTDGGLYLWQPLIQNLRLDKN--KL 435
Db 340 FKCIIEQVN-KTHEKYEAVQDRYTKGOEALTYFTTSGLLALLMLPLTPRSLATVKNLTTEL 398
Qy 436 NNETYS-----RRSRQAESTTDPMMEMTGNGAGEYSSSENSI 473
Db 399 TPTSPSPSPSPAPSAARGSTPAAVLRRRRRDAGNATP---VPPTAPGKSLGTLLNP 455
Qy 474 TVAQVOVAYDNLRIRINNILEDLSKAWCRQHRRAALVMNELSKINPTSVMSMIYNRPVSA 533
Db 456 ATVQIQFAYDSLRRQINRMGLDLARAWCLEQKQNMVLRLTKINPTTVMSSYIGKAVAA 515
Qy 534 KRIGDVISVNCIIVDQTSVSLHKSURLLSASDEKCFSPRPVTFKFMNDSTIYKQGLGVN 593
Db 516 KRLGDVISVQCPVNAQATVTLRKSMT-VPGSETMCSRPLVSFSFINDTKTYEGQLGTD 574
Qy 594 NEILLTTTLETQENTYFYFOAKTDMYIKYNEHLKTVPLSSITTLDTIALNFTLLEN 653
Db 575 NEIFLTKKTEVCQATSQYFQSGNEIHVYNDYHHPFKTIELDGIATLQTTISLNTSLIEN 634
Qy 654 VDFKVELYTRDEKRLSNVFDIETMPREYNYYAQRYSGLRKDLDD-LSTNRNQFVDAFGS 712
Db 635 IDFASLELYSRDEQRASNVFDLEGI FREYNFQANIAGLRKLDLNAVSNCRNPFVDGLGE 694
Qy 713 LMDDLGAVGQTVVNAVSGVATLFFSSIVTGFINKPFGGMLMIIVVIGLVFAIYFLTKX 772
Db 695 LMDSLGSGQITNLVSTVGLFSSLVSGFISFFKNPFGMLILVLVAGVILVSLTR 754
Qy 773 TKIVETAPIKMIYPEIDKLKEREKSE---IAPISEEELERIVLAMIHOONSHMETKTR 829
Db 755 TROMSQOPVQMLYPGIDELAQQHASGEGPGINPISKTELQAIMLA--LHEQNOEQKRAAQ 812
Qy 830 KDPKDSILTRAQNMRLKR 847
Db 813 RAAGPSVASRALQAARDR 830

RESULT 40
US-10-267-748-103
; Sequence 103, Application US/10267748
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; Matthews, Thomas J.
; Wild, Carl T.
; Barney, Shawn O.
; Lambert, Dennis M.
; Petteway, Stephen R.
; Langlois, Alphonse J.

; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
; MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
; TRANSMISSION
; NUMBER OF SEQUENCES: 239
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA

575	NEIFLT	KMTVE	COATS	QY	YFQSG	NIHVND	VHHKFT	ITEL	GIAT	FLQT	FI	SLNT	SLIN	633
654	VDFKVI	LYTR	DEKRL	SNV	FDI	ETM	FR	YNYA	QR	VS	GLR	KOLL	D	LS
635	IDFAS	LELY	S	DE	Q	RAS	NV	FD	LE	GI	FR	Y	N	FO
713	LMDD	LGA	V	Q	T	V	N	AV	S	G	V	AT	L	P
695	LMDS	L	G	S	V	G	S	I	T	N	L	V	S	T
773	TKIV	ET	API	K	W	I	P	E	I	D	K	L	R	E
755	TR	MS	Q	P	V	Q	M	L	P	G	I	D	E	L
830	KDP	K	S	I	L	T	R	A	Q	N	M	L	R	K
813	RA	A	G	P	S	V	A	S	A	L	Q	A	R	D
<p>RESULT 41</p> <p>US-08-804-439-17</p> <p>Sequence 17, Application US/08804439</p> <p>GENERAL INFORMATION:</p> <p>APPLICANT: Rose, Timothy M.</p> <p>APPLICANT: Bosch, Marnix L.</p> <p>TITLE OF INVENTION: GLYCOPROTEIN B OF THE RPHV/KSHV</p> <p>TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES</p> <p>NUMBER OF SEQUENCES: 113</p> <p>CORRESPONDENCE ADDRESS:</p> <p>ADDRESSEE: Mortison & Foerster</p> <p>STREET: 755 Page Mill Road</p> <p>CITY: Palo Alto</p> <p>STATE: CA</p> <p>COUNTRY: USA</p> <p>ZIP: 94304-1018</p> <p>COMPUTER READABLE FORM:</p> <p>MEDIUM TYPE: Floppy disk</p> <p>COMPUTER: IBM PC compatible</p> <p>OPERATING SYSTEM: PC-DOS/MS-DOS</p> <p>SOFTWARE: Patentin Release #1.0, Version #1.30</p> <p>CURRENT APPLICATION DATA:</p> <p>APPLICATION NUMBER: US/08/804,439</p> <p>FILING DATE:</p> <p>CLASSIFICATION: 424</p> <p>ATTORNEY/AGENT INFORMATION:</p> <p>NAME: Schiff, J. Michael</p> <p>REGISTRATION NUMBER: 40,253</p> <p>REFERENCE/DOCKET NUMBER: 29938-20002.20</p> <p>TELECOMMUNICATION INFORMATION:</p> <p>TELEPHONE: (415) 813-5600</p> <p>TELEFAX: (415) 494-0792</p> <p>TELEX: 706141</p> <p>INFORMATION FOR SEQ ID NO: 17:</p> <p>SEQUENCE CHARACTERISTICS:</p> <p>LENGTH: 849 amino acids</p> <p>TYPE: amino acid</p> <p>STRANDEDNESS: single</p> <p>TOPOLOGY: linear</p> <p>MOLECULE TYPE: protein</p> <p>US-08-804-439-17</p>														
<p>Query Match</p> <p>Best Local Similarity 46.1%; Pred. No. 1.5e-163;</p> <p>Matches 368; Conservative 148; Mismatches 240; Indels 42; Gaps 13;</p>														
Qy	81	STP	YV	CS	AS	G	V	D	V	F	R	Q	T	H
Db	44	T	S	P	F	R	C	L	S	S	G	D	L	F
Qy	140	T	T	S	V	N	G	I	Y	S	D	I	T	N
Db	104	T	N	I	L	I	N	G	W	A	D	S	V	T
Qy	200	M	T	V	L	P	O	V	D	G	V	T	P	D
Db	164	I	T	V	N	L	K	P	T	G	L	A	N	G
Qy	260	T	A	T	G	D	T	V	E	M	S	P	F	F
Db	224	T	T	T	G	T	V	E	M	S	P	F	F	D
Qy	320	Y	T	L	S	W	E	K	H	K	N	-	M	S
Db	282	Y	T	L	S	W	K	-	L	E	N	R	T	A
Qy	379	Y	H	C	L	N	E	I	K	A	E	I	K	K
Db	340	F	K	C	I	E	E	Q	V	N	-	K	T	H
Qy	436	N	N	E	T	I	S	-	-	-	-	-	-	-
Db	399	T	T	T	S	P	S	P	S	P	P	A	A	G
Qy	474	T	V	A	Q	V	A	Y	D	N	L	R	I	N
Db	456	A	T	V	O	I	Q	A	I	S	L	R	Q	N
Qy	534	K	R	I	G	D	V	I	S	N	C	I	W	I
Db	516	K	R	L	G	D	V	I	S	Q	C	V	P	N
Qy	594	N	E	I	L	L	T	T	T	L	E	T	C	O

Db 668 YSLAGIKKDLNDNTIDYNRDLVQDLSDMMADLGDIGRSVVNVSSVTFSSIVTGFIKF 727
Qy 746 IKNPFGGMLIIWIGVFAIYFLTKTKIYETAPIKMIYPIBIDKLKREG----KSBIA 801
Db 728 FTNPLGGIFILLIIGIIFLVVLRNRSQFHDAPIKMLYPSVENYARQAPPPYSASPP 787
Qy 802 PISEELERIVLAMH-IHQONSHMETK-TRKDPKDSILTRAQNMRLR-KRSGYSNL 853
Db 788 AIDKEIKRILLGMHVHQBEEKAQKLTNSGP--TLWQKATGFLNRNRKGYSQL 840

RESULT 43
US-09-338-326-17
; Sequence 17, Application US/09338326
; GENERAL INFORMATION:
; APPLICANT: Rose, Timothy M.
; APPLICANT: Bosch, Marnix L.
; APPLICANT: Strand, Kurt
; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RFHV/KSHV
; TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/338.326
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/720,229
; FILING DATE: 26-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Schiff, J. Michael
; REGISTRATION NUMBER: 40,253
; REFERENCE/DOCKET NUMBER: 29938-20002.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 849 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-338-326-17

Query Match 39.4%; Score 1792; DB 17; Length 849;
Best Local Similarity 42.4%; Pred. No. 4.1e-159;
Matches 354; Conservative 158; Mismatches 271; Indels 52; Gaps 13;

Qy 34 SPPNTATWSTPLTGHYGTHTSDSHGERGNENRDESEQNKNIGSPSTFPYRVCSAGV 93
Db 43 TPAQAPTEPTPLS-----TWNRCFEY-----FRVCGVAAT 75
Qy 94 GDVFRQTDHVCDDASDMVHSEGILLIYKONIIPFMFRVRYKRVKVVTTSTVNGIYSDSI 153
Db 76 GETFRDLKTCPSODKKHVEGILLVYKINIVPIFKIRYKRIITQITWRGLTTSV 135
Qy 154 TNOHTFYKSIPEWETEKMDTYQCNSRLNTGGNLLTYVDRDDINMTVFLQPDGVTPD 213
Db 136 TCKFEMATQAEHWEVGDPSIYQYNSATMVNVNRYVYVDRDGVNKTNIRPVDGLTGN 195

Qy 214 VKRYGSOPELYLEPCWFGWSYRRRTTNCLELMDMFARSNPDPDFVVTATGDTVMSPFWS 273
Db 196 IQRYFSQPTLYSEPGWMPGFYVRVTTCVCEIVDMVARSDPYNVIATATGLSLELSPQT 255
Qy 274 GEDHENKMHK-PWFVSVINNYKVYQNRGTVPGLKTRIFPLDREEVTLSEKHLKNMS 332
Db 256 FDNTSQCTAPKRAMRVREKNYKFDVYNNRGTPAGOSRTFLETPTSATYSWKATROTA 315
Qy 333 YCPILTLMKAFYNGTQTEHSGSYHFVANDITASFTTSKEDMKEPNTTYHCLNEEIKAEK 392
Db 316 TCDLVHMKTFPRAIQTAEHSYHFVANEVATFNTPLTEVENFTSTYSCVSDQINKTISE 375
Qy 393 KYAKVNSTHSGYDKYFKTDGGLYLVWQPLIQNLRLDANKLNNE-----TYSRRSRQ 447
Db 376 YIQLNNSYVASGKTQYFKTDGGLYLVWQPLHEPEIEDIDSDPEPTPAPPKSTRKRE 435
Qy 448 AESTTDPMMETGAGAGEYSSSENSITVAQVQAYDNLIRINILLEDLSKAWCREQURA 507
Db 436 AADNGNSTSEVS-----KGSNPLITAQIQFAYDKLTTSVNVNVEELSRAWCREQVRD 488
Qy 508 ALVNELSKINPTSMNRYNRPVSAKRGDVISVNCIVVDOTSVSLHLSLRLLSASDE 567
Db 489 TLMYELSKVNPTSVNGAIYKPVAAVYGVDAISVTDICIVDQSSVNIHQSLR-LQHDKT 547
Qy 568 KCFSRPPVTFKFMNDSTIYKQGLGVNNEILLTTLVLETCQENTYVYFOAKTDMVIYKNYE 627
Db 548 TCYSRPRVTFKFINSTDPGLTQGLPRKEIILSNINIECKDESEHYFVIGEVYIYKNYI 607
Qy 628 HLKTVPLSSITTLDTFIALNFTLLENVDFKVIELYTRDEKRL-SNVFDIETMFPREYNYYA 686
Db 608 FEELNLSSITATLDTFIALNISFIENIDFKVELYSSTERKLASSVFDIESMFREYNYT 667
Qy 687 QRVSLRKDLDD-LSTNRNQPVDAPGSLMDLGAQGVQTVNNAVSGVATLSSIVTGFINF 745
Db 668 YSLAGIKKDLNDNTIDYNRDLVQDLSDMMADLGDIGRSVVNVSSVTFSSIVTGFIKF 727
Qy 746 IKNPFGGMLIIWIGVFAIYFLTKTKIYETAPIKMIYPIBIDKLKREG----KSBIA 801
Db 728 FTNPLGGIFILLIIGIIFLVVLRNRSQFHDAPIKMLYPSVENYARQAPPPYSASPP 787
Qy 802 PISEELERIVLAMH-IHQONSHMETK-TRKDPKDSILTRAQNMRLR-KRSGYSNL 853
Db 788 AIDKEIKRILLGMHVHQBEEKAQKLTNSGP--TLWQKATGFLNRNRKGYSQL 840

RESULT 44.
US-08-804-439-15
; Sequence 15, Application US/08804439
; GENERAL INFORMATION:
; APPLICANT: Rose, Timothy M.
; APPLICANT: Bosch, Marnix L.
; APPLICANT: Strand, Kurt
; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RFHV/KSHV
; TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,439
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Schiff, J. Michael

QY	273	SGEDDHENKM-----HEKPMFVSVNNYKVVDYQNRGTVPLGKTRIFLDRREYTL	322
Db	283	-----HNSQCTGNSTRDATKFW---IEENHQTVDYRRGH-PTKDRIEFLKDBEYTI	333
QY	323	SWEXHLKNM5CPLTLWKAFYNGIOTEHSGSVHYFVANDITASFTTSKEDMKEF--NT-TY	379
Db	334	SWKAEDRERACDFVIWKTFPRAIOTIHNE5PHFVANEVTAGFLTSNQETSELRGNTEIL	393
QY	380	HCLNEEIKAEIEKKYAKVNSTHSKYGDLKYFKTDGGLVLWQOPLIQNRLLDKRN---KLN	436
Db	394	NCMNSTINETLEETVKFKNKSHIRDEGVKYYKTNGGLFLIWQAMKPLNLSEHTNVTIERN	453
QY	437	NETYSRRSRQAE5TDPMEMTNGAGGEYSSENSITVAQVOYAYDNLRIINNILEDL	496
Db	454	NKTGNKSRQKRSVDPT-----KTFOGAKG-----LSTAQVOYAYDHLRTSMNHILBEL	500
QY	497	SKAWCREQHRAALVMNELSKINPT5VMGMIYNRPVSAKRIGDVISV5NCIIVDDQTSVSLH	556
Db	501	TKTWCREQKONLMWYELSKINPVSVMALYKGPVAVKAWGDAFWVSECINVDQASVNIH	560
QY	557	KSLRLLSASDEK-CFSRPPVTFKFNDSTIYKGQGVNNEILLTTTYYLETQCENTEYYPQ	615
Db	561	KSMR---TDDPKVCYSRPLVTFKPVNSTATFRGQLGTREIILLTNTHVETCRPTADHPFF	617
QY	616	AKTDMYIYKNYEHKATVP5LSTTLTDTFLANFTLLENVDKVIELYTRDEKLSNVPTI	675
Db	618	VKNTHYFKDFYFKVKTMDTNISTUDTFLTLMTFIDNIDFKTVEL5ETEKRMASALDL	677
QY	676	ETMPREYNYYAQRV5GLRKDL---LDLSTNRNQFVDAFGSLMDDELGAGQGTVVNAV5GVA	732
Db	678	ETMPREYNYYQKLASREDLDNTIDL--NRDLVKDLSEMMADLGDIGKVVVNTFSGIV	735
QY	733	TLF5SIVTGFNFINKNPGGMLMIIVLGVLPAYIYFLTKTKIYETAPIKMIYPIDKUK	792
Db	736	TVFSGIVGGFVSFFNPTIGGVTIILLLLVVPWFIVSRRTNNMNEAPIKMIYPNIDKAS	795
QY	793	EREKGE5IAP5EBELERIVLAMHTHQNSMHTKTRKDPKD---SILTRAQNMRLKESG	849
Db	796	EQE---NIQPLPG5EIKRLLGMHQLQOSEHGK5EEASHKPGLPQLLDGDLQLLURRR-G	851
QY	850	YSNL	853
Db	852	YTRL	855

Search completed: October 28, 2003, 15:39:45
Job time : 403 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 15:30:15 ; Search time 31 Seconds
(without alignments)
1276.284 Million cell updates/sec

Title: US-10-055-364-24
Perfect score: 4547
Sequence: 1 MAGSLKRGSVLALWLYQV.....KRGYSNLKNAESVEMLNTL 865

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 242630 seqs, 45739658 residues

Total number of hits satisfying chosen parameters: 242630

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents_AA_New.*

- 1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep.*
- 2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep.*
- 3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep.*
- 4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep.*
- 5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep.*
- 6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep.*
- 7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1813	39.9	854	5	US-09-350-841A-1589
2	844	18.6	904	1	PCT-US03-11231-18
3	836.5	18.4	904	1	PCT-US03-30301-56
4	789.5	17.4	943	6	US-10-670-695-14
5	244.5	5.4	227	1	PCT-US03-30301-54
6	148.5	3.3	1431	6	US-10-679-063-18001
7	135.5	3.0	1638	1	PCT-US03-26635-1
8	130.5	2.9	774	6	US-10-425-114A-43519
9	130.5	2.9	1276	6	US-10-205-516A-8
10	128	2.8	1001	1	PCT-US03-20460-10
11	128	2.8	1193	6	US-10-679-063-18223
12	126.5	2.8	3899	1	PCT-US02-18638A-4
13	126.5	2.8	3907	1	PCT-US02-18638A-2
14	126.5	2.8	3917	1	PCT-US02-18638A-8
15	126.5	2.8	3925	1	PCT-US02-18638A-6
16	126	2.8	9121	6	US-10-679-063-18203
17	124.5	2.7	353	6	US-10-425-114A-55825
18	124	2.7	250	6	US-10-670-695-16
19	124	2.7	1038	1	PCT-US03-27401-460
20	124	2.7	1038	6	US-10-472-928-4532
21	123	2.7	526	6	US-10-472-928-3638
22	122.5	2.7	1286	6	US-10-205-516A-22
23	122	2.7	1664	1	PCT-US02-24459-102
24	122	2.7	1738	1	PCT-US02-24459-100
25	121.5	2.7	836	6	US-10-679-063-11928
26	121	2.7	1147	6	US-10-679-063-18214

27	121	2.7	2004	1	PCT-US03-27401-364	Sequence 364, App
28	121	2.7	2004	6	US-10-472-928-2306	Sequence 2306, App
29	120.5	2.7	846	6	US-10-679-063-18007	Sequence 18007, A
30	120	2.6	793	6	US-10-679-063-18197	Sequence 18197, A
31	120	2.6	1732	6	US-10-679-366-2	Sequence 2, Appli
32	118.5	2.6	1280	1	PCT-US03-27401-274	Sequence 274, App
33	118.5	2.6	1280	6	US-10-472-928-364	Sequence 364, App
34	117.5	2.6	928	1	PCT-US03-30720-1209	Sequence 1209, App
35	117.5	2.6	1030	6	US-10-425-114A-62748	Sequence 62748, A
36	117	2.6	880	6	US-10-330-773-956	Sequence 956, App
37	117	2.6	889	6	US-10-679-063-18166	Sequence 18166, A
38	116	2.6	1301	6	US-10-205-516A-18	Sequence 18, Appl
39	116	2.6	1881	1	PCT-US03-27401-316	Sequence 316, App
40	116	2.6	1881	6	US-10-472-928-1236	Sequence 1236, A
41	116	2.6	2497	6	US-10-679-063-18067	Sequence 18067, A
42	114.5	2.5	1169	6	US-10-472-928-4712	Sequence 4712, App
43	113.5	2.5	1431	6	US-10-679-063-20989	Sequence 20989, A
44	113.5	2.5	1518	1	PCT-US03-06962-40	Sequence 40, Appl
45	113.5	2.5	1617	6	US-10-679-063-18152	Sequence 18152, A

ALIGNMENTS

RESULT 1

US-09-350-841A-1589
; Sequence 1589, Application US/09350841A
; GENERAL INFORMATION:
; APPLICANT: Jeffs, Peter;
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION OF MEMBRANE
; FILE REFERENCE: 7872-066-999
; CURRENT APPLICATION NUMBER: US/09/350.841A
; CURRENT FILING DATE: 1999-07-09
; NUMBER OF SEQ ID NOS: 1946
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1589
; LENGTH: 854
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-09-350-841A-1589

Query Match	39.9%	Score 1813;	DB 5;	Length 854;
Best Local Similarity	46.0%	Pred. No. 8.4e-91;		
Matches	367;	Conservative 147;	Mismatches 240;	Indels 44; Gaps 15;
Qy	81	STFPYRVCSAGVGDFRFTQDHVCPD-ASDMVHSEGLLIYKQNIIPFMRFRVKYKV 139		
Db	44	TSPFPRVCELSSHGDLFRFSSDIQCPSGFTRENHTEGLLMVFKDNIIPYSEKRSYTKIV 103		
Qy	140	TTSTVYNGIYSDSITNOHTFYKSTIEPWETEKMDTIYQCFNSLRNLNTGNNLLTYVDRDIN 199		
Db	104	TNLIYNGWYADSVNTRHEEKFSVDSYETQMDTIYQYNAVKMTKDLTRVYVDRDGVN 163		
Qy	200	MTVELQPVGVTPDVVKYGSOPELYLPFGWFGSYRRRTTVCNCELMDFARSNPPDFPV 259		
Db	164	ITVNLKFTGGLANGVRVYASQTELYDAPGLMIWTYRTTVCNCLITDMAKSNPPDFPV 223		
Qy	260	TATGDTVEMGPFWSGEDDHENKMKHEKPFWSVSVNNYKVQYQNRGTVPGLKTRIFLDREE 319		
Db	224	TTTQQTVMSPFFYDGK--NKETFERADSFHVRTNYKIVDYNRGTNPQGERAFDKGT 281		
Qy	320	YTLSEWHLKN-MSYCPDLTLWAFYNGIOIEHSSYHFVANDITASTTTSKEDKENTT 378		
Db	282	YTLSSWK--LENRTAYCPLQHWOTFDSITATETGKSHFVTDGTSSTFVNTTVGIELPDA 339		
Qy	379	YHCLNEBIKAEIEKAKVNSVTHSKYGD-LKYFKTDGGLVLMQPLIQNLIDAKN--KL 435		
Db	340	FKCIEEQVN-KTHKYEAVQDRYTKGQEAITYFITSGLLLALWLPITPRSLATVKNLTTEL 398		
Qy	436	NNETYS-----RRSRROAESTTDPMMEMTNGAGGEGYSSNSI 473		
Db	399	TTPTSSPPSPSPAPSAARGSTPAVLRRRRRDAGNATTP---VPPTAPGKSLGTLLNP 455		

QY	474	TVAQOVAYDNLRIRINILEDLSKAWCREQHRAALVNNELSKINPTSVMSIYNRPVSA	533
Db	456	ATVQIQAYDSLRQINRMGLDARAWCLEQKQNWVLRILTINKIPTVWSSYIGKAVAA	515
QY	534	KRIGDVISVNCIIVDQTSVLSLHKSRLLSASDEKCFSPRPVTFKFMNDSTIYKGLGVN	593
Db	516	KRLGDVISVQCVPVQATV-LRKSME-VPGSETMVCYRPLVSFSPFINDTKTYEGQGTD	573
QY	594	NEILLITTYLETQENTYEFQAKTDMYIYKNEVHLKTVPLSITTLDTFIALNFILLEN	653
Db	574	NEIFLTKMTEVCQATSYQYFQSGNEIHVYNDYHFKTIELDGIATLQTPIS-NTSLIEN	632
QY	654	VDFKVELYTRDBKRLSNVPDIETMREYNYAQRVSGRLKDLDD-LSTNRQFVDAFGS	712
Db	633	IDFASLELYRDRQRANSVFDLGEIFREYNFQACNIAGLRKDJDNVNGNRNQFVDGLGE	692
QY	713	LMDDLGAVGQTVNAVSGVATLFESSIVTGFINFKNPFGGMLMIIVVIGVLFAYIFLTKK	772
Db	693	LMDSLGSVGQSIITNLVSTVGGLFSSLSVSGFISFFKNPFGGMLILVLVAGVILVISLTR	752
QY	773	TKIYETAPIKWIYPIEDIKLEREKGSE---IAPISBEELERIVLAMIHOONSHMETKTR	829
Db	753	TRQMSQOPVQMLYFGIDELAQHASGEGPGINPISKTELQAIMLA--LHEQNOEQKRAAQ	810
QY	830	KDPKDSILTTRAQNMRLKR	847
Db	811	RAAGPSVASRALQARDR	828
RESULT 2			
PCT-US03-11231-18			
; Sequence 18, Application PC/TUS0311231			
; GENERAL INFORMATION:			
; APPLICANT: Corixa Corporation			
; APPLICANT: Day, Craig H.			
; APPLICANT: Hosken, Nancy A.			
; APPLICANT: Parsons, Joseph M.			
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND			
; TITLE OF INVENTION: TREATMENT OF HERPES SIMPLEX VIRUS INFECTION			
; FILE REFERENCE: 210121.53801PC			
; CURRENT APPLICATION NUMBER: PCT/US03/11231			
; CURRENT FILING DATE: 2003-04-09			
; NUMBER OF SEQ ID NOS: 267			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 18			
; TYPE: PRT			
; ORGANISM: Herpes simplex virus			
PCT-US03-11231-18			
Query Match 18.6%; Score 844; DB 1; Length 904;			
Best Local Similarity 28.4%; Pred. No. 1.7e-38;			
Matches 247; Conservative 164; Mismatches 366; Indels 92; Gaps 27;			
QY	33	TSPNTATWSTESPLTGHYTHDSSHGGERGNENRDSSEQKNLYGSPSTFPYRVCSASG	92
Db	76	TPPPD-----ANATVAAGHATLR--AHUREIKVENADAQ-----FYVCPPT	115
QY	93	VGVDFRQTDHVCPSDASD-MVHSEGIILYKQNIIPFMRVRKYRKVVTSTV-----	144
Db	116	GATVQFEQPRCPTRPREGQNYTEGIAVFPKENIAPYKFKATMYKDVTSQVWFGRHYS	175
QY	145	-YNGIYSDSTINQHTFYKSTIEPWE--TEKMDTYQCNSLRNTGNNLLTYVDRDDINMT	201
Db	176	QFMGIFED-----RAPVPEEVIDKINTKVCRCSTAKYVRNNMETTAFHRDDHETD	226
QY	202	VELQPDVGVTPDKRYGSQLYLEPCGFWGVSRRRTTVNCELMDWFARSNPDPDEPVA	261
Db	227	MELKPAK-VATRTSRGWHITDLKYNPSRVEAFHRYGTVVNCIVVEEDARSVYYPDEPFLA	285
QY	262	TGDTVMSPEWS--GEDDH--ENKMHEKFWFSVINNY--KVVDYQNRGVTVPLGKTRIFLDRE	318
PCT-US03-30301-56			
; Sequence 56, Application PC/TUS0330301			
; GENERAL INFORMATION:			
; APPLICANT: SYKES, KATHRYN F.			
; APPLICANT: HALE, KATHERINE S.			
; APPLICANT: JOHNSTON, STEPHEN A.			
; TITLE OF INVENTION: METHODS FOR VACCINE IDENTIFICATION AND COMPOSITIONS FOR			
; TITLE OF INVENTION: VACCINATION COMPRISING NUCLEIC ACID AND/OR POLYPEPTIDE			
; FILE REFERENCE: MCRO:002WO			
; CURRENT APPLICATION NUMBER: PCT/US03/30301			
; CURRENT FILING DATE: 2003-09-23			
; PRIOR APPLICATION NUMBER: 60/412,956			
; PRIOR FILING DATE: 2002-09-23			
; NUMBER OF SEQ ID NOS: 116			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 56			
; LENGTH: 904			
; TYPE: PRT			
; ORGANISM: Herpes Virus			
PCT-US03-30301-56			
Query Match 18.4%; Score 836.5; DB 1; Length 904;			
Best Local Similarity 27.5%; Pred. No. 4.3e-38;			
Matches 244; Conservative 181; Mismatches 361; Indels 101; Gaps 28;			
QY	24	SLSIATGTVPNTATWSTESP--LTGHYTHDSSHGE--RGNENRDSSEQKNLYGSP	80
Db	58	ALGAAPTGPKPKKKNKPKNPTPPRAGDNATVAAGHATLRHLRDLKAENTDANFY---	114

```
QY 81 STFPYRVCSASGVDFRQTDHVCPCDASD-MVHSEGLIYKQNIIPFMFRVRYKRV 139
Db 115 -----VCPPTGATVVOEQPRCPTRPEGQNYTEGIAVFKENIAPVKFKATWYKDV 168
QY 140 TTSTV-----YNGIYSDSITNOHTFYKSTIEPWE--TEKMDTTIQCFSNLRNLNGNL 189
Db 169 TVSQWFMGHRYSQFMGIFED-----RAPVFEVIDKINAKGVCRCSTAKYVRNNLE 219
QY 190 LTYVDRDDINMTVFLQPDGV-----TPDKRYGSOPELYLBPFGWFSYRRRTTVN 241
Db 220 TTAHRRDDHETDMLKPAWAATRTSRGHTTDLKKNPSRVEAF-----HRYGTTVN 270
QY 242 CELMDAFARNPPDFPFVATGDTVMSPFWS-GBDDH-ENKMHKPMFVSINNY-KV 298
Db 271 CIVEEDARSVVPYDFVLATGDFVVMSPFYREGSHTEHTSYAADRKFQVDGYARDL 330
QY 299 DYONRGTVPLGTRIFLDREBYTLSEKHLKMSVCPITLWKAIFYNGIQTEHSGSYHVA 358
Db 331 TTAKATAP--TTRNLLTTPKFTVAMDVWPKRPSVCTMTKQEVDEMRLSEYGGSF 388
QY 359 NDITASFT-----SKEDMKFNTTYHCLNEEIKAEIEKKYA-KVNSTHSKYGDLKYF 410
Db 389 DAISTFTTNLTPELPSRVLDG-----CIKDARDAMDRIFFARYNATHIKVQPOQY 442
QY 411 KTDGGLYLWQPLIQNLRLDAKNKNNETYSRRSROAESTTDDPMEMTNGAGGEYSSE 470
Db 443 LANGGFLIAYQPLLSNTLAE-----LYVREHLREQSRKPNPTPPPGASANASVERIKTT 498
QY 471 NSITVAQVQYADNRIIRINNILEDLSKAWCREQRAALVNLKSNPTSVMSMYNRP 530
Db 499 SSIEFARLOFTYTHIQRHVNMDLGRVAIAWCELOHNEHLTLNNEARKLPNAIASATVGR 558
QY 531 VSAKRIGDVISVNCIVVDQTSVLSHLSRLSASDEKCFSPPTVTFKEMDSTIYKGL 590
Db 559 VSARMIGDVMAVSTCPVAADNVIVONSNR-ISSRPGACYSRPLVSFRYDQGPLVEGOL 617
QY 591 GYNNEILLTITTYLETQCENTEYFOAKTDMYIKNYEHLKTVPLSSITTLDTFIALNFTL 650
Db 618 GENNELRLTRDAIEPCTVGHRRYFTFGGYVYFESAVSHQSLRADITTVTFIDLNTM 677
QY 651 LENVDKUIELYTRDEKLSNVFDIETMFREYNYAQRVSLGRLKDLLDLSTNRNQF--VD 708
Db 678 LEDHEFVPLEVYTRHEIKDGLLDYTEVQRNQLHDLRFADI-DTVIHADANAAMAFALG 736
QY 709 AFGSLMDLGL-AVGOTVNAVSGVATLPSIIVTGFINFIKPNFGMLIIVIGVLFAY 767
Db 737 AFFEGMDLGRAVGKVMGVIIGGVV-----SAVSGVSMSPNPFGLAVGLLVLAGLAAAF 792
QY 768 FLTKTKIYETAPIKMIYPEIDK-LK-----EREKSEIAPISE-EELERIVLA 814
Db 793 FAFRYVMRLQSNPKALYPLITTKELKNPTNPDSAGEGEGDFDEAKLAAREMIRYNAL 852
QY 815 MHIHQONSHMETKTRKPKDSILT-RAQNML---KRSGYSNLKNAE 857
Db 853 VSAMERTHE---KAKKKGTSALLSAKVTDVMVRKRRTNYTQVPNKD 896
```

RESULT 4

```
US-10-670-695-14
; Sequence 14, Application US/10670695
; GENERAL INFORMATION:
; APPLICANT: Jensen, Wayne A.
; APPLICANT: Lappin, Michael R.
; APPLICANT: Rosen, David K.
; APPLICANT: Andrews, Janet S.
; TITLE OF INVENTION: USE OF RECOMBINANT ANTIGENS TO DETERMINE THE IMMUNE
; STATUS OF AN ANIMAL
; FILE REFERENCE: DI-9-1
; CURRENT APPLICATION NUMBER: US/10/670,695
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: 09/521,738
; PRIOR FILING DATE: 2000-03-09
```

```
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 14
; LENGTH: 943
; TYPE: PRT
; ORGANISM: Feline herpesvirus 1
US-10-670-695-14
```

```
Query Match 17.4%; Score 789.5; DB 6; Length 943;
Best Local Similarity 27.3%; Pred. No. 1.5e-35;
Matches 231; Conservative 151; Mismatches 364; Indels 101; Gaps 27;

QY 80 PSTFPYRVCSASGVDFRQTDHVCPCDASD-MVHSEGLIYKQNIIPFMFRVRYKRV 138
Db 121 PSTP--YMCPPSGSTVRLPEPRACPDYKLGKFTGEGIAVIFKENIAPYKFKANIYKN 178
QY 139 VTTSTVNGIYSDSITNOHTFYKSTIEPWE--TEKMDTTIQCFSNLRNLNGNLTYVDRDD 197
Db 179 IIMTVMSGSYAVTTNRYTDRVPVKQEIITDLIDRGMCLSKADYVRNNYQFTAFDRDE 238
QY 198 INMTVFLQPDGVDPDKRYGSOPELYLBPFGWFS-----YRRRTTVNCLMDM 247
Db 239 DPRELPLKPSSTLSRV-----GWHTNETHYTKIVLDDFHSGHSTVNCIVEEV 286
QY 248 FARNPPDFPFVATGDTVMSPFWSGED-----DHENKMHKPMFVSINNYKVVDYQNR 303
Db 287 DARSVVPDSFAISTGDVIMHSPFFGLRDGAHVEHTSYSSDR---FQIEGYYPIDLDTD 343
QY 304 GT-VPLGKTRIFLDREBYTLSEKHLKMSVCPITLWKAIFYNGIQTEHSGSYHVFANDIT 362
Db 344 YTGAPV--SRNPLETPHTVAVNMTPKSGRVCTLAKWREI--DEMPLMNIGSYRFTAKTIS 400
QY 363 ASFTTSKEDMKFNTTY--HCLNEEIKAEIEKKY-KVNSTHSKYGDLKYPKTGGLYLV 419
Db 401 ATFTISNTSQF-EINRIKLGDCATKEAABADRIYKSKYKTHIOTGTLETYLARGGFLIA 459
QY 420 WQPLIQNL-----LDANKLN--NETYSRRSR-----RQESTTDP 454
Db 460 FRPMISNELAKYINELARSNRTVVDLSALNPSEIVQTRRSVPSNQHRSSRTIEG 519
QY 455 MMENTGAGGEYSSENSITVAQVAYDNRIIRINNILEDLSKAWCREQRAALVNL 514
Db 520 GIETVNNAS--LLKTTSSVEFAMLQFAYDIQAHVNEMLSRITATAWCTLQNRHVLWMTET 577
QY 515 SKINPTSVMSMYNRPVSRAKIGDVISVNCIVVDQTSVLSHLSRLSASDEKCFSPRP 574
Db 578 LKLNPGGVSNALERRVSARLLGDVAGVATQCVNISSGHVYIIONSMR-VTGSSTTCYSRPL 636
QY 575 VTFKPMNDSTIYKQGLGVNNEILLTITTYLETQCENTEYFOAKTDMYIKNYEHLKTVPL 634
Db 637 VSFRALNDSEYIEQOLGENNELVERKLI.EPCTVNNKRYPKFGADYVYFEDYAVRVKPL 696
QY 635 SSITTLDTFIALNTLLENVDKUIELYTRDEKLSNVFDIETMFREYNYAQRVSLGRLK 694
Db 697 SEIELISAYV-IKSTLLEDREF-LHSSYTRAELDGTGPFYDSEIQRNQLHALKFYDI-- 752
QY 695 DLLDLSTNRNQFVDAFGSLMDLGLAVGQTVNAVSGVATLPSIIVTGFINFIKPNFGML 754
Db 753 DSVIRVDNNLVIMRGMANFFQGLGDVGAGFGKVLGAASAVISTVSGVSSPLNPFPGALA 812
QY 755 MIIWIGVLFAYELTKTKIYETAPIKMIYPEIDKLEREGKSE-----IAP 803
Db 813 VGLLILAGIYVAAFLAYRYISRLRANPKALYPVTRNLKQTAKEPASTAGSDSPGVDDF 872
QY 804 SEELER-----IVLAMHIHQONSHMETKTRKPKDSILT-RAQNMLRKRSK--YS 851
Db 873 DEELMQAREMIKYMVSLVSAM---EQOEHKAMKKNKGP--AILTSLTNMALRRRGPKYQ 927
QY 852 NLKNAES 858
Db 928 RLNNLDS 934
```

RESULT 5

PCT-US03-30301-54
 ; Sequence 54, Application PC/TUS0330301
 ; GENERAL INFORMATION:
 ; APPLICANT: SYKES, KATHRYN F.
 ; APPLICANT: JOHNSTON, STEPHEN A.
 ; TITLE OF INVENTION: METHODS FOR VACCINE IDENTIFICATION AND COMPOSITIONS FOR
 ; TITLE OF INVENTION: VACCINATION COMPRISING NUCLEIC ACID AND/OR POLYPEPTIDE
 ; TITLE OF INVENTION: SEQUENCES OF THE HERPESVIRUS FAMILY
 ; FILE REFERENCE: MCR0:002W0
 ; CURRENT APPLICATION NUMBER: PCT/US03/30301
 ; CURRENT FILING DATE: 2003-09-23
 ; PRIOR APPLICATION NUMBER: 60/412,956
 ; PRIOR FILING DATE: 2002-09-23
 ; NUMBER OF SEQ ID NOS: 116
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID-NO 54
 ; LENGTH: 227
 ; TYPE: PrT
 ; ORGANISM: Herpes Virus
 PCT-US03-30301-54

Query Match 5.4%; Score 244.5; DB 1; Length 227;
 Best Local Similarity 28.3%; Pred. No. 1.1e-06;
 Matches 67; Conservative 52; Mismatches 95; Indels 23; Gaps 8;

Qy 254 PFDFVTATGTVMSDFWS-GEEDH-ENKMEKPFVSVINNY-KVVDYQNRGTVPGLGK 310
 Db 3 PYDEFVLATGDFVMSPFYGYREGSHTSHTSYAADREFKQVDGFYARDLTTKARATAP-T 60
 Qy 311 TRIFLDREEVTLSEKHLKMSYCPLTLWKAIFYNGIOTHSYVHFVANDITASFTT--- 367
 Db 61 TRNLTLTPKFTVADWVPKPSVCTMTKQOEVDMLRSEYGGFRFSSDAISTFTTNLT 120
 Qy 368 ---SKEDMKFNTYHCLNEEKAEIEKKYA-KVNSTHSGYKGLYFKTDGGLYLWQP 422
 Db 121 EYPLSRVDLGD-----CIGKARDAMDRIFARRYNATHIKVQPOYYLANGGFLIAYQP 174
 Qy 423 LIONRLDANKLNNTYRSRRQAESTTDPMMETGNGAGGEYSSENITVAQVQ 479
 Db 175 LLSNTLAE---LYVREHLREQSRKPPNTPPPPGASANASVERIKTSSIEFARLQ 227

RESULT 6

US-10-679-063-18001
 ; Sequence 18001, Application US/10679063
 ; GENERAL INFORMATION:
 ; APPLICANT: Edgerton, Michael D
 ; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
 ; FILE REFERENCE: 38-15(52054)B
 ; CURRENT APPLICATION NUMBER: US/10/679,063
 ; CURRENT FILING DATE: 2003-10-02
 ; PRIOR APPLICATION NUMBER: 60/415,758
 ; PRIOR FILING DATE: 2002-10-02
 ; NUMBER OF SEQ ID NOS: 27373
 ; SEQ ID NO 18001
 ; LENGTH: 1431
 ; TYPE: PrT
 ; ORGANISM: Plasmodium yoelii yoelii
 US-10-679-063-18001

Query Match 3.3%; Score 148.5; DB 6; Length 1431;
 Best Local Similarity 18.2%; Pred. No. 0.86;
 Matches 192; Conservative 154; Mismatches 338; Indels 373; Gaps 55;

Qy 50 HYGTHDSHGERGNENR-----DSEQNKNIYVGPSTF 83
 Db 242 NYDNNTNVRDNNQSNYSYADENETKDNNEEDDDTGDYDDNEEDDDNDNDYS 301
 Qy 84 PYRVCSASGVDGVRFF-----QTDHVCPSADSMVHSEGIL-----LIY-KONI 125
 Db 302 QXNQCEVESDPTNQIRSNRKYNESIKHI-NESNLKINKESLLKRETYNKRDNIFYIKDOI 360

Qy 126 IPE-----MFRVRKYRKVTTSTVYNGIYSDSITN-OHTFYKSIETPMT-----EK 170
 Db 361 IPYKKEHNNNIFSLYDSSK---NNNEHNNNYETKFVNYKHGDEKKDEQMLTNVKKERNEE 417
 Qy 171 MDTIYQCFNSL-----RLMTGGNLLTYVDRDDINMTVFVLQPDVG-----TPDV 214
 Db 418 FLKKYMYENALKTYSSKDLQFNRLGDEKNILHH-----DINVDNKMFRDLKLSRNTODI 473
 Qy 215 KRYGSOPELYLEBPGWPGWGYRRRTT-----VNCELMDMFARSNPPDFVVTATGDTVM 268
 Db 474 RNKTNK-----YNNFRHDTTSLSKIRYKDEINDYSEYANKCLE-----GNIYD- 516
 Qy 269 SPWMSGEDD-----HENKMEKPFVSVINNYKVVDYQNRGTVPGLGKTRI--FL 315
 Db 517 -----EDDSYLKRKELGHKELKVIDNP-----ILN-----INHDDK-----KKTNIKSF 557
 Qy 316 DREYTLSEKHLKN-----MSYCPLTLWKAIFYNGIOTHSYVHFV 357
 Db 558 DKIKYRKNNELFLKNEHEKYADIQETNKKDKIMKEADILHTKNFSNF-----HRGSSIA 614
 Qy 358 A-----NDITASFTT-----SKEDMKF-----NTVH-CLNEIK- 387
 Db 615 KSLSPNNKKNANFNNEAFNLNILEKKDKNSDYKQYNIKCFPYDHENTSPHPSVRBETKY 674
 Qy 388 -----AEIEKKYAKVNSTHSGYKGLYFKTDGGLYLWQPLIQNRLD 430
 Db 675 YENKESRKLVDYCDENDIENSIRYNENNIDYNSINETIN-----LREDIKYNEFLN 730
 Qy 431 AKNKLNNETYRR---SRQAESTTDPMMETGNGAGGEYSSENS-----ITVAQVQYAY 482
 Db 731 KPDKINSEKFNNSFNDSNKKSLKNDDSNKIRDTNKLYYHNSDNALRSNVIKKESTEKIY 790
 Qy 483 DNLIRINNILEDLSKA-WCROHRAALVWNE---LSKIN--PTSVMYMIYNRPVSAKRI 536
 Db 791 DILESNNMRKFNKMDVYTKYEQNLKNDEDIYTKENKFPDS-ROAFYKNSNSPMKV 849
 Qy 537 GCVISVS-----NCIVV-----DQTSVLSLHSLRLLLSASDEKCFSPPTTFKF 579
 Db 850 KOETKDSQVEDYNRKIYKUKMKQEKODEDETDINIDKKLSRKMIENNK-----LKE 901
 Qy 580 MNDSTYKQGLGVNNEILLTIVLET-----COBENTYV-----FOAKTDMYIKN 625
 Db 902 QNES-----NDELIVTPFYKSKIDKVLKNSEIFERSARATFKQFDVKN-----KN 947
 Qy 626 YHLKTVPLSSITTLDTFIALNFTLLENVDFKVIEL-----YTRDEKRLSNVFIETMFE 681
 Db 948 FLH-----FSEIESLTKLCYNLE-LPPVDKKILSIYVKDYDSSKKNCMYMDFROMYWD 1001
 Qy 682 Y-----NYVAQVSGRLKDLDDLSTNRNQVDVAFGLSMDLLGAVGQTVVNAVSGVATLF 735
 Db 1002 LKQIKKKYPTKNFKIKRNCI-ISRKKLQGYD-----Y 1034
 Qy 736 SSIVTGFNFIK-----NPFGGMLMIITWIGVLFAIYFLTKTKTK-----774
 Db 1035 SSIY-NYLSFKILGCGAFGEVHLVEDNICKLYKVVKILKKKKMKNIKVNEEINVLILYD 1093
 Qy 775 -----IYETAPIKMIYPEI-----DKLKEREKSEIAPISEEBLERIVLAWHH 4818
 Db 1094 HPNIKIFDVYESVNCYIWMELCEGELMSKIKPEIFSEIY-IKNIMFOILCAIAYMH 1152
 Qy 819 QQN-SHMETKTRKDPKDSILTIRAQNMRLKRSYGNLK 854
 Db 1153 SNIAHKDLK-----PENILFKTDGYDTLK 1177

RESULT 7

PCT-US03-26635-1
 ; Sequence 1, Application PC/TUS0326635
 ; GENERAL INFORMATION:
 ; APPLICANT: INCYTE CORPORATION; BAUGHN, Mariah R.;
 ; APPLICANT: RICHARDSON, Thomas W.; WARQUIS, Joseph P.;
 ; APPLICANT: SWARNAKAR, Anita; Tang, Y. Tom;

; APPLICANT: BECHA, Shanya D.; EMERLING, Brooke M.;
 ; APPLICANT: JIN, Pei; WILSON, Amy D.;
 ; APPLICANT: YUE, Henry; GIETZEN, Kimberly J.;
 ; APPLICANT: CHANG, Hsin-Ru; YANG, Yonghong G.;
 ; APPLICANT: LEE, Soo Yeun; KHARE, Reena;
 ; APPLICANT: ELLIOTT, Vicki S.; HAPALIA, April J.A.;
 ; APPLICANT: CHAMLA, Navinder K.; RAMKUMAR, Jayalaxmi;
 ; APPLICANT: GURAJARAN, Rajagopal; TRIBOULEY, Catherine M.;
 ; APPLICANT: CHIEN, David; TRAN, Uyen K.;
 ; APPLICANT: MURAGE, Jaji
 ; TITLE OF INVENTION: KINASES AND PHOSPHATASES
 ; FILE REFERENCE: PF-1566 PCT
 ; CURRENT APPLICATION NUMBER: PCT/US03/26635
 ; PRIOR FILING DATE: 2003-08-25
 ; PRIOR APPLICATION NUMBER: US 60/406,172
 ; PRIOR FILING DATE: 2002-08-26
 ; PRIOR APPLICATION NUMBER: US 60/413,910
 ; PRIOR FILING DATE: 2002-09-25
 ; PRIOR APPLICATION NUMBER: US 60/414,296
 ; PRIOR FILING DATE: 2002-09-27
 ; PRIOR APPLICATION NUMBER: US 60/417,821
 ; PRIOR FILING DATE: 2002-10-11
 ; NUMBER OF SEQ ID NOS: 118
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 1
 ; LENGTH: 1638
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; OTHER INFORMATION: Incyte ID No: 7504529CD1
 PCT-US03-26635-1

Query Match 3.0%; Score 135.5; DB 1; Length 1638;
 Best Local Similarity 18.6%; Pred. No. 4.9;
 Matches 124; Conservative 95; Mismatches 201; Indels 247; Gaps 31;
 QY 151 DSITNQHTFYKSIETWETEKMDTIYQCENS-URLNTGGNLLTYVDRDDINMTVFLQPDVG 209
 DB 190 DSVHQHLYVHRDIKP-----DNILMDXNGHRLADFSCLXLMEDGTIVQSSVAVGTPDY 243
 QY 210 VTPDV-----KRYGSOPE-----LYLEPGWF-----WG-----SYRRRTTV 240
 DB 244 ISPEILQAMEDGKGRYGPEDCWMSLGVCMYEMLYGETPFYABSLVETYGKIMNHRERQF 303
 QY 241 NCELMDMFARSNPPDFVTATGDTVEMS-----PWSCEDDHENKMHKPFVSV- 291
 DB 304 PAQVTDVSENADPIRRLLICGHEHRLGSGIEDFKKHPFPGSIDWDNIRNCEAPYIPEVS 363
 QY 292 ----INNYKVVD--YONRGTVP-----LQGT 311
 DB 364 SPDTNSFVDVDDCLKNSTMPPTHTAPSGHHLFPVGFTYTSSCVLSDRSCLRVTAQPT 423
 QY 312 RI-----FLDREYTLSEKHLKNMSYCPLLTWKAFYNGIQTGSHGSYHFVANDITA 363
 DB 424 SLDDLVNVQRTLDNNLATEAYERRIKRLEQKLESLRKLQESTQIVQALQYSTVDGPLTA 483
 QY 364 SFTTSKEDMKENTYTHCLNEBI--KAELEKKYAKVNSTHSHYGLD-----KYFTD 413
 DB 484 SKDLEIKNLKE--VIEKLRKVQTESHLEQOOLEEAVRQELDDAFRQIKAYEKQIKT- 539
 QY 414 GGLYLWQPLIQNRLDAKNLNNETYS-----RRSRROAESTTDDPMEMTNGAGG 465
 DB 540 -----LQER--EDLNKLEVITEALAAASKDKLREQE----- 572
 QY 466 EYSS--ENSIT-VAQVQVAYDNLIRINNILEDLSKAMCREQHRALVWNLKINPTSV 522
 DB 573 HYSKOLENEGLKQKQISY-----SPGVCSIEHQ-----QETTKLKTDL 613
 QY 523 MSMIYRNPVSAKRIGDVTISVNCIVVDQTSVSLH-----KSLRLLSASDEKCFSPVPTFK 578
 DB 614 KKSIFYEELSREG-----IHANEIKNLK-----K 639

QY 579 FMNDSTIYKG-QLGWVNEIL-----LTTTYLETQENTYFYFOAKTDWYIYKNYEHLKTV 632
 DB 640 ELHDS---EGQALNKEIMILKDKLEKTRRESQSREPESEFK-----QYEREKV- 689
 QY 633 PLSSITTLDTFIALNFTLLENVDFKVIELYTRDKRLSNVFD-IETMPREYNYVAQRVSG 691
 DB 690 -----LSTEENKLTSELDKLTTLTYENLSIHNOOLEE 721
 QY 692 LRKDLDD 698
 DB 722 EVKDLAD 728
 RESULT 8
 US-10-425-114A-43519
 ; Sequence 43519, Application US/10425114A
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E
 ; APPLICANT: Tabaska, Jack E
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
 ; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53313)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114A
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 43519
 ; LENGTH: 774
 ; TYPE: PRT
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: 700468168_FLI.pep
 US-10-425-114A-43519

Query Match 2.9%; Score 130.5; DB 6; Length 774;
 Best Local Similarity 19.8%; Pred. No. 4.6;
 Matches 101; Conservative 78; Mismatches 203; Indels 129; Gaps 20;
 QY 434 KLNNE-TYSRRSRQAESTTDPMMEMTNGAGGEY--SSENSITVAQVQVAYDNLIRIN 490
 DB 30 KTNELAINFNLSSANTTINSLODALSQARSNIFILAAEKNEAEKAYETEINALNAELT 89
 QY 491 NILEDLSKAMCREQHRALVWNLKIN-----PTSVNMI---YRNPVSAKRIGDVTISVS 543
 DB 90 KCLBELDKTHGLQSHSTHYGYLEKLGFMFMDDSLLSLAEYEGKTFNSLR-----D 142
 QY 544 NCIVVDQTSVSLHKSRLLSASDEKCFSPVPTFKMNDSTIYKGL----- 590
 DB 143 MCLIVK-----SMHQLSVKS-----FONDSIVEDSELSSLPLPDYESFV 183
 QY 591 -----GVNNEILLTTTYLETQENTYFYFOAKTDWYIYKNYE---HLKTVPLSS 636
 DB 184 RELVKAINRKNIGDITSSFSSTIVEQLSNQTEYLSFLKSLSTYMSNIIIVLRLSLQVS 243
 QY 637 ITTLDTFIALNFTLLENVDFKVIELYTRD-----EKRLSNVFDIETMPREYNYVAQRV 689
 DB 244 NTFART-----LEEDHMLKVELGNKDAHRAQAESEVLSQLDKRAMSSKCIYCVQOI 295
 QY 690 SGLRKDLDD-----STNRNQVDAFGSLMDDLGAVGQTVVNAV 728
 DB 296 EIVFDDMVDLGYAIDLATGSSGISELVTVDLKNEDTSDYSKVAOTLTLATIDLKSKS 355
 QY 729 SGVATLFPSSIVTGFNF---IKNPFGG-----MLMIIVWIGVLFAIYFLTKTKTIYE 777
 DB 356 EKLSTIKGLVITSLDDDFKMLKQAEAAEAETASHEHQLSV-----ERVCMLEKELIKLQ 408
 QY 778 TAPTKMTYPIDKLKEREGKSEIAPISEBELERI-VLAMHIHQONSHMETKTRKDKPKDSI 836
 DB 409 DECNMEL-NIQEYKEREK-----ALKARELELLSVHEHTQISADRLGTDNAISKQDEAL 462

Db 1038 VNKSGEESKIMVEDKVSFENMTVGEESKQEQILDLHLPSTVKSSLRATQPSNDKLOK 1097
Qy 787 EIDKLEREGKSEIAPISBEELERIVLAM-----HIHQONSHMETKTRK 830
Db 1098 ELNVKSEQNDLRL-----QMEAQRICLSLVYSTHVDQVREYMENEDK 1141

RESULT 13

PCT-US02-18638A-2
; Sequence 2, Application PC/TUS0218638A
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc. et al.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035PC
; CURRENT APPLICATION NUMBER: PCT/US02/18638A
; PRIOR FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 3907
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-18638A-2

Query Match 2.8%; Score 126.5; DB 1; Length 3907;
Best Local Similarity 18.7%; Pred. No. 33;
Matches 110; Conservative 101; Mismatches 196; Indels 181; Gaps 26;
Qy 373 KEFTTYHCLNEEIKAEIEKKYAKVNSTHSGYDLYKFTDGLGYLVWQPLIQNRLLDK 432
Db 605 KKNVLDRAESEAELERLRTQLFSHEE--ELSKLED-----LEIE 647
Qy 433 NKLNNETYSRRRQAEISTTDPMMMTGNGAGGY-----SSENSITVAQVQYVDNL 485
Db 648 HRIN-----IEKLKDLGHIHYKQIDGLQNMESQKTIETWQFEDK 688
Qy 486 RIRNNILEDLSKAWCREQ-----HRAALVWNLSEK-----INPT 520
Db 689 ITKQNLILEISKDLQOGLVNSKSEEMTLQINELQKEIEILRQEEKGTLEQEQVEL 748
Qy 521 SVMSIMYNNRPVSASRIGDV-----ISVNCIVVDQ-----TSVSLHKSRLLS 563
Db 749 QLKTELLEKQMKKE--NDLQEKFAQLEAENSILKDEKKTLEDMLKIHTPVSQERLIFLD 807
Qy 564 ASDEKCFSPPTFK-----FMNDSTIYKGO--LGVNNEILLTTTYLETQCENTEYFQ--- 615
Db 808 SIKSK--SKDSVWEKEIEILIEENEDLKQOCIQLENEIEKQRTNTPFAEKNFVNYQELQ 865
Qy 616 -----AKTDMYIKNYEHLK-----TVPL--SSITLTD--TFIA 645
Db 866 EYACLLKVKDDLESDSKNQKELEYKSLKALNEELHLQRLNPPTVMKSSVFPDEKTFVA 925
Qy 646 LNFTLLENVDFKVEL-----YTRDEK-----RLSNVFD--IETMFREYNYAQRVSLGRK 694
Db 926 ETELEMGVEVEKDTTELMEKLEVTREKLELSQRLSDLSQKQKHGEISFLNEEVSKLKQ 985
Qy 695 DL-----LDLSTNRNQFVDAFGLMDDLGAQGTQVNVNAGVATLFPSSIVTGFNF 745
Db 986 EKEQVSLRCREGEIINHNH-----AENVQSCDTQVSSLLDGVVVTMTSRGAEGSVSK 1037
Qy 746 IKNPFGGMLMIIV-----VIG-----VLFPAIFLTTKTKIYETAPIK--MIYP 786
Db 1038 VNKSGEESKIMVEDKVSFENMTVGEESKQEQILDLHLPSTVKSSLRATQPSNDKLOK 1097

Qy 787 EIDKLEREGKSEIAPISBEELERIVLAM-----HIHQONSHMETKTRK 830
Db 1098 ELNVKSEQNDLRL-----QMEAQRICLSLVYSTHVDQVREYMENEDK 1141
RESULT 14
PCT-US02-18638A-8
; Sequence 8, Application PC/TUS0218638A
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc. et al.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035PC
; CURRENT APPLICATION NUMBER: PCT/US02/18638A
; PRIOR FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 3917
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-18638A-8

Query Match 2.8%; Score 126.5; DB 1; Length 3917;
Best Local Similarity 18.7%; Pred. No. 33;
Matches 110; Conservative 101; Mismatches 196; Indels 181; Gaps 26;
Qy 373 KEFTTYHCLNEEIKAEIEKKYAKVNSTHSGYDLYKFTDGLGYLVWQPLIQNRLLDK 432
Db 605 KKNVLDRAESEAELERLRTQLFSHEE--ELSKLED-----LEIE 647
Qy 433 NKLNNETYSRRRQAEISTTDPMMMTGNGAGGY-----SSENSITVAQVQYVDNL 485
Db 648 HRIN-----IEKLKDLGHIHYKQIDGLQNMESQKTIETWQFEDK 688
Qy 486 RIRNNILEDLSKAWCREQ-----HRAALVWNLSEK-----INPT 520
Db 689 ITKQNLILEISKDLQOGLVNSKSEEMTLQINELQKEIEILRQEEKGTLEQEQVEL 748
Qy 521 SVMSIMYNNRPVSASRIGDV-----ISVNCIVVDQ-----TSVSLHKSRLLS 563
Db 749 QLKTELLEKQMKKE--NDLQEKFAQLEAENSILKDEKKTLEDMLKIHTPVSQERLIFLD 807
Qy 564 ASDEKCFSPPTFK-----FMNDSTIYKGO--LGVNNEILLTTTYLETQCENTEYFQ--- 615
Db 808 SIKSK--SKDSVWEKEIEILIEENEDLKQOCIQLENEIEKQRTNTPFAEKNFVNYQELQ 865
Qy 616 -----AKTDMYIKNYEHLK-----TVPL--SSITLTD--TFIA 645
Db 866 EYACLLKVKDDLESDSKNQKELEYKSLKALNEELHLQRLNPPTVMKSSVFPDEKTFVA 925
Qy 646 LNFTLLENVDFKVEL-----YTRDEK-----RLSNVFD--IETMFREYNYAQRVSLGRK 694
Db 926 ETELEMGVEVEKDTTELMEKLEVTREKLELSQRLSDLSQKQKHGEISFLNEEVSKLKQ 985
Qy 695 DL-----LDLSTNRNQFVDAFGLMDDLGAQGTQVNVNAGVATLFPSSIVTGFNF 745
Db 986 EKEQVSLRCREGEIINHNH-----AENVQSCDTQVSSLLDGVVVTMTSRGAEGSVSK 1037
Qy 746 IKNPFGGMLMIIV-----VIG-----VLFPAIFLTTKTKIYETAPIK--MIYP 786
Db 1038 VNKSGEESKIMVEDKVSFENMTVGEESKQEQILDLHLPSTVKSSLRATQPSNDKLOK 1097
Qy 787 EIDKLEREGKSEIAPISBEELERIVLAM-----HIHQONSHMETKTRK 830
Db 1098 ELNVKSEQNDLRL-----QMEAQRICLSLVYSTHVDQVREYMENEDK 1141

APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114A
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 55825
LENGTH: 353
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: UC-GMROPIC049A02_FLI pep
US-10-425-114A-55825

Query Match 2.7%; Score 124.5; DB 6; Length 353;
Best Local Similarity 20.2%; Pred. No. 4.7;
Matches 80; Conservative 65; Mismatches 143; Indels 109; Gaps 19;
QY 455 MMENTGAGGEVSSSENSITVAQVQYAYDNLIRININILEDLSKAWCREQHRALVWNL 514
Db 30 LFEELGOSAGDPLVEEDDIPILRSWOSQNFPLVTWMI-----KGSVSW--I 74
QY 515 SKINPTSVMSMI-----YNRPVSAKRIGDVISVSNCIIVVDQTSVSLHKSRLLSASDEKCF 570
Db 75 NVLIGITEVQELLSAGGYNMP---RTVHEVIALACRLSRWDDRLFRAS--IFGNADE--- 126
QY 571 SRPVTFKFN-----DSTYKGGOLGVNNEILLTTTYLETQCENTEYVFOAKTDMYIYKN 625
Db 127 -----IELKFNRRNHEDLNLF--ILLNQEI---RKLSTQVIRVKSLSHAR-DEIVFEL 175
QY 626 YEHLKTVPLSITLDTFIALNFTLLENVDFKVIETREKLSNVDFDIETFREYNY 685
Db 176 LQHLKGNAR-----TLLEGIKKSTREIEEQEAVRGLFTIQVNMQ----- 217
QY 686 AQRVSGLRKOLLDLSTNRNQFVDAGSLMDDLGAVGQTVVNAVSGVATLFSSIVTGFINF 745
Db 218 ----STVRAMLODRSLRVTHNLAVF-----GGVG-VLTIITG----LFGINVDGIPGA 262
QY 746 IKNP--FGMLMIIVIGVLFVFAIYFLTKIYETAPIKMIYPIIDKLKEREGKSEIAPI 803
Db 263 EQTPYAFGVFTAILVWLGVL-----IAGVMVYLGKN-----PV 297
QY 804 SEELE-----RVLAMHHIQNSHMETKTRDKPKD 834
Db 298 ABEQVEVRKUELOELVKNQFOHEAETHAQMRRKNISPKN 334

RESULT 18
US-10-670-695-16
Sequence 16, Application US/10670695
GENERAL INFORMATION:
APPLICANT: Jensen, Wayne A.
APPLICANT: Lappin, Michael R.
APPLICANT: Rosen, David K.
APPLICANT: Andrews, Janet S.
TITLE OF INVENTION: USE OF RECOMBINANT ANTIGENS TO DETERMINE THE IMMUNE
FILE REFERENCE: DI-9-1
CURRENT APPLICATION NUMBER: US/10/670,695
CURRENT FILING DATE: 2003-09-25
PRIOR APPLICATION NUMBER: 09/521,738
PRIOR FILING DATE: 2000-03-09
NUMBER OF SEQ ID NOS: 36
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 16
LENGTH: 250
TYPE: PRT
ORGANISM: Feline herpesvirus 1
US-10-670-695-16

Query Match 2.7%; Score 124; DB 6; Length 250;
Best Local Similarity 28.7%; Pred. No. 3.7;
Matches 37; Conservative 22; Mismatches 66; Indels 4; Gaps 3;
QY 80 PSTFPYRVCSASGVGVDFQTDHVCPDAS--DMVHSEGILLIYKQNIIPFMFRVRYKRV 138
Db 121 PSTF--YMCPPPSGSTVVRLEPPRACPDYKLGKNTFEGIAVIFKENIAPYKFKANIYKN 178
QY 139 VTTSTVNGIYSDSITNOHTFYKSIPEWE--TEKMDTIYOCFNSLRNLNTGNNLLTYVDRDD 197
Db 179 IMTIVWSSGYAVTTNRYTDRVPVKVQEIITDLIDRRGCLSKADYVRNNYQTFAPDRDE 238
QY 198 INMTVFLQP 206
Db 239 DPRELPLKP 247
RESULT 19
PCT-US03-27401-460
Sequence 460, Application PC/TUS0327401
GENERAL INFORMATION:
APPLICANT: TUFTS UNIVERSITY
TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE FOR DIAGNOSIS, TREATMENT AND
FILE REFERENCE: 700355-52941-PCT
CURRENT APPLICATION NUMBER: PCT/US03/27401
CURRENT FILING DATE: 2003-09-02
PRIOR APPLICATION NUMBER: US 60/407,082
PRIOR FILING DATE: 2002-08-30
NUMBER OF SEQ ID NOS: 560
SOFTWARE: Patent in version 3.2
SEQ ID NO 460
LENGTH: 1038
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
PCT-US03-27401-460

Query Match 2.7%; Score 124; DB 1; Length 1038;
Best Local Similarity 17.8%; Pred. No. 14;
Matches 120; Conservative 84; Mismatches 216; Indels 254; Gaps 27;
QY 209 GVTDPVKRYGSDPELYLEPGWFGWGYRRRTTNNVNCCLMDMPARNPFPDF-----FVITATGD 264
Db 267 GRADMRSYASEPE-----SMIAMENMNVYTGCTVYVNFECAYTFMTND 311
QY 265 -----TVMSPEWSGEDDHENKHEKWPWFVSVINNYKVVDYQNRGTVPLGKTRIFLDRE 318
Db 312 VPTPAFTKGIIPFRHAIQNPAPSKEE-----VVRNRTKAVFMNGEGRI----- 354
QY 319 EYTLSEKHLKNMSYCPLTLWKAFYNGIQTE-----HSGSYHFV-----ANDITA 363
Db 355 -----SSLNGFYQGLYSNDETMTPLYNNGRYHILPVIHEKIDKEKISS 396
QY 364 SFTTSKEDMKEFNTTYHCLNEEIKABIEKKYAKVNSTHSGYDLKFKTDDGGLYLVWQPL 423
Db 397 IFPNKILTKN-----SEELS-----SKVNYLSLYP--KLYEGDG----- 430
QY 424 IQNELLDKKNLNNETYSRSRQAEASTTDPMMEMTNGAGGYSSENSITVAQVOYAYD 483
Db 431 -----YAQRVGNWYIYNSNANINKNQOVMPLPMTNTKSLSLDTPHYAVVKNPN 483
QY 484 NLRIRINILEDLSKAWCREQHRALVWNLSPINPTSVMSIYNNRPSAKRIGDVISVS 543
Db 484 NLHILLNNYRTDKTAMW-----ALSGNFDASKSKKK--EELELA 520
QY 544 NCIVDQTSVSLHKSRLLSASDEKCFSRPPTVTFKMNDSITYKQGLGVNNEILLT----- 599
Db 521 NWI-----SKNYSINPVNDPRTTLLTKGHTGHPQINISGDKN 560
QY 600 -TTYLETQCENTEY-----FOAKTDMYIK-----NYEHLKTVPL 634
Db 561 HYTYTENNDENTHVTITVNHNGMVMSINTEGTPVSPFPTPKFNDGNLNIAYAKTTQ 620

```

635 SSI-----TTLDTFIALNF-----TLLENVDFKVIELYTRDEKLNSNVF---DIET 677 Qy
   |||      :||       |         |         |         |         |         |
621 SSVDYGDNPDRADGNRRNGENSGSVTHRADNPSWWEVLKKMDKVGLVKIYNRTDAET 680 Db
   |||      :||       |         |         |         |         |         |
678 MREINYYAQRVSIGRKDLLDLSTNRNQFDAGFSIMDDLGAVGQTVVNANSGVATLFPSS 737 Qy
   |||      :||       |         |         |         |         |         |
681 -----QLSNF--DVILYDNNRNE-----VAKHVVNLISGESVSLDF 715 Db
           |||      :||       |         |         |         |         |
738 IVTGFINFKNPFGGMLIIIVVIGLVFAIFYLTKTKIIVETAPIKMVIPEIDKLKEREGK 797 Qy
          |||      :||       |         |         |         |         |
716 KEKG-----ARYI----VKULL-TSGVPLSIAEVEVFRESOGK 748 Db
                           |||      :||       |         |         |         |
798 SEIAPISEELERI 811 Qy
        |||      :||       |         |         |         |         |
749 Q-----SEEDIDKI 757 Db


RESULT 20
US-10-472-928-4532
; Sequence 4532, Application US/10472928
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; TITLE OF INVENTION: THE INSTITUTE FOR GENOMIC RESEARCH
; FILE REFERENCE: P026926WO
; CURRENT APPLICATION NUMBER: US/10/472,928
; PRIOR FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: GB-0107658.7
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 4979
; SOFTWARE: Seqwin99, version 1.03
; SEQ ID NO 4532
; LENGTH: 1038
; TYPE: PRM
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; OTHER INFORMATION: fucollectin-related protein
; OTHER INFORMATION: Cellular location: outside
; OTHER INFORMATION: Similar to strain R6 sequence 15904006 (O.E+01)
US-10-472-928-4532
```

Query Match	2.7%;	Score 124;	DB 6;	Length 1038;
Best Local Similarity	17.8%;	Pred. No. 14;		
Matches	120;	Conservative	84;	Mismatches 216;
				Indels 254;
				Gaps 27;
QY	209	GVTPDVKRYGSOPELYLEPGFWGCSYRRRTIVNCELDMFARSNPPDF	----	FVTATGD 264
DB	267	GRADMSYASEP		
QY	265	-----TVEMSPFWSGEDDHENKHKHPFVSVINNYKVVDYQNRGTVPLGKTRIFLDRE	----	318
DB	312	VPTPAFTKGIIPFPHRAIQNPAPSKEE		----
QY	319	EYTLSEWKHLKNM5CYPLTLWKAFYNGIQTE	-----	HSGSYHFV-----ANDITA 363
DB	355	-----SSLNGYQGLYSNDETMPLYNNGRYHILPVIHEKIDKEKISS	----	396
QY	364	SFTTSKEDMKEFNTYTHCLNBEIEKAEIKYAKVNSTH5KYGDLYKFTDGLGLVWQPL	423	
DB	397	IFPNAKILTKN-----SEELS-----SKVYLSNLYP--KLYEGDG-----	430	
QY	424	IQNELLDKNNKLNETYSRSRROAESTTDDMMEMTNGAGGESSENSITVAOVQYAYD	483	
DB	431	-----YAQRVGNSWYIYSNAINKNQOQVMLPMYTNNTKSLDLTLPHFYAVVKNPN	483	
QY	484	NLRIRINNILEDLSKAWCREQHRAAALVWNL5KINPTSVMSMIYNNRPVSAKRIGDVISVS	543	
DB	484	NLHILLNNYRDKTAMW-----ALSGNFDASKSWK--EELELA 520		
QY	544	NCIVVDQTSVLHLSKRLLSASDEKCPSPRPVTFKFNNDSTIYKGOLGVNNEILLT	----	599
DB	521	NWI-----SKNYSINPVDNDFRTTTLTKGHTCHKPQINISGDKN	560	

```

QY      600   -TTYLETQENTEYY-----FQAKTDMVIYK-----NYEHLKTVLPL 638
           ||| ||||| :| |
Db      561   HYTYTENWENDVTITVNHNGVMVMSINTEGTGSPVSFTPKDKNCGNGLNIAYAKPTTQ 620
QY      635   SSI-----TTLTOTFALNF-----TLLENVDKFVIELYTRDEKRLSNVF---DIET 677
           ||| :| | ||| :| |
Db      621   SSVDYNGDPNRADVGNRNGNSGVSVTHTRADNPSSWEVDLKMDKVGLVKIYNRTDAET 680
QY      678   MFREYNYAQRVSGLRKDLLDLSTNRNQFVDAFGSLMDDLGAVGOTVVNAVSGVATLFSS 737
Db      681   -----QLSNSP--DVILIYDNRRNE-----VAKHVNLSGESVSULDF 715
QY      738   IVTGFINFIKNPFGMLMIIVIGVLFAFYFLTKTKIYETAPIKMIYPEIDKLKEREG 797
Db      716   KEKG-----ARYI---KVKLL-TSGVPLSLAEVEVFRESDBG 748
QY      798   SEIAPISSEELERI 811
Db      749   Q-----SEEDIDKI 757

RESULT 21
US-10-472-928-3638
; Sequence 3638, Application US/10472928
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026926WO
; CURRENT APPLICATION NUMBER: US/10/472,928
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: GB-0107658.7
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 4979
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 3638
; LENGTH: 526
; TYPE: PRP
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; OTHER INFORMATION: hypothetical protein
; OTHER INFORMATION: Cellular location: cytoplasm
US-10-472-928-3638

Query Match          2.7%; Score 123; DB 6; Length 526;
Best Local Similarity 18.2%; Pred. No. 8.2; Indels 126; Gaps 27
Matches 108; Conservative 107; Mismatches 252;

QY      270   PFWSGEDDHENKMHE--KPWFVSVINNYKVVDYQNRGTVPLGKTRIF---LDREEYTLIS 323
           ||| ||| :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| |
Db      6     PAWYGS---ERTWHADITPWYFSHR---LEFDD---TFHQIRLFOEQDIDSRLLVLA 54
QY      324   WEKHLKNMSYCPLTLWKAFYNGIQTEHSGSYHFVANDITASFTTSKEDMKGFNTYTHCLN 383
           ||| :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| |
Db      55   YQPHRYFLRYRHGHVLEMDTYSVFDMQ--DFPNLHTQVLSIRDIEWDDDCEFYSPFTI- 111
QY      384   BEIKABIEKKYAKNSTHSHKYG---DLKYFKTDGGLYLWQPLQNRLLDANKLNNETY 440
           ||| :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| |
Db      112   --IVQKNKGKPAKVE--HGVEGFISDIQYFPNGOIHW-----HHLVDRGRFVSSIIF 160
QY      441   SRRSRRQAESITDPMEMETGN-----GAGEYSSENSITVAQYQAYNDLRIRINNIL 493
           ||| :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| |
Db      161   F---EDQAAQOEYLNKGEWQFRERLKEGGQEVNVPILGYRFQMILTQNMGDIVAEREFF 216
QY      494   EDLSKAWCREQ-----HRAALVNNELSKINPTSVMSIMYINRPVSAKRIGDV-ISVS 543
           ||| :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| |
Db      217   ENYLQTYYVKQODIEMLPSSHHDQLVLDRLPSTNPKLLSLFIGRNPDQTFRDLDVTFPEKS 276
QY      544   NCIVVDQTSVSLHKSLRLLSASDEKFSRPPVTFKFWNDSTIYKLGOLGVNNELLTTTYVL 603
           ||| :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| | :| |
Db      277   DLILVDR-----EDSLRL-----QSLYPERMHQCYHLSSFDTRL 311

604   ETCENTHEYFOAKTDMVIYKYNVEHLKTVPLSSITTLDTFALNPTLLENVDFKVI-ELY 662

```

Db 312 RLGRS-----QTKESIYFQDFEQIDNQALQVLFSFA-----ENKDTVEIFCAF 359
Qy 663 TRDEKRLSNVFDIETMFREYNYAQRVSGRLKDLDDST--NRNQFVD---AFGSLMDDL 717
Db 360 AASQEQMNEVEGIESFTQENIQSEN--GKAIDYGDADENPLEENQHQDLRLQFVNLDEL 418
Qy 718 GAVG-----QTVNAVSGVATLFSSIVTGFIFINFKPFGMLMIWVI 760
Db 419 DLIKTLEFVRLIVDLNRPHLYTOIAGISAGIPQINLVETVYVBLKNGY----LLADVT 474
Qy 761 GVLFALYELTKTKIYETAPIKMIYPEIDKLKREGKSEIAPIS--BELERI 811
Db 475 EFSKAAHYTDBLKEWNS---LIY-SIDKIKEHTGQOFLGKLEKWEIEVKV 523

RESULT 22

US-10-205-516A-22
; Sequence 22, Application US/10205516A
; GENERAL INFORMATION:
; APPLICANT: Zhong, Jun
; TITLE OF INVENTION: A Method For Producing Biologically Active Botulinum Neurotoxins
; FILE REFERENCE: JZbtcl
; CURRENT APPLICATION NUMBER: US/10/205,516A
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 22
; LENGTH: 1286
; TYPE: PRT
; ORGANISM: Clostridium botulinum
US-10-205-516A-22

Query Match 2.7%; Score 122.5; DB 6; Length 1286;
Best Local Similarity 18.8%; Pred. No. 20;
Matches 147; Conservative 118; Mismatches 274; Indels 243; Gaps 41;
Qy 64 NENDRSEONKNIYGSSTFPRVCSAGVCDVFRFQTDHVC-----PDASDMV 112
Db 403 NKGFIENSQONIERNFA---LQKLSSESVVDLF---TKVCLRLTKIEGRNSRDDSTCI 455
Qy 113 HSEGIILY---KNIITPFMPVRKRYKVVTTST--VYNGIYSDISITNQHTFYKSIPEWT 168
Db 456 KVKNRILPYVADKDSISQEIFE-----NKIITETNVQN--YSDKFSIDESILDCQVPI 509
Qy 169 EKMDTIYQCFNSRLNTGGNLLTYVDRDDINMTVFLQPDVGTDPVKRYGSQPELYLEPG 228
Db 510 EIVDPLLPNVNMEPLNLPGBEIVFYD-----DITKYVD---YLSY 547
Qy 229 WFWGYSR-----RTTVNCELMDMFARSNPPDF---FVATGDTVMSPP--WSGE--D 276
Db 548 YLESQKLSNNVENITLTSVEEALGYSNKIYTFELPSLAEKVKVGQAGFLNANVEVE 607
Qy 277 DHENKMEKHPWF-----VSVNNY---KVVDYQNRG-----T 305
Db 608 DFTTNIMKOTLDKISDVSIIPGIPALNIGNSALRGNFQAPATAGVAFLEGPFEFT 667
Qy 306 VP-LGKTRIFLDREYTLSEKHLKNSYC---PLTLWKAFYNGIOTEHSGSYHFVANDI 361
Db 668 IPALGVFTFYSSIQER---EKIKTIENCLQORVKWK-----DSYQMVSNW 712
Qy 362 TASFTTKEDMKEN---TTHCLN---BEIKAEIEKKYAKV-----N 398
Db 713 LSRITT-----QFNHINYQMYDSLSYQADAIAKAKIDLEYKRYKSGDKENIKSQVENLKN 766
Qy 399 STHSKYGD-----KYFKTDGGLVWQPLIQNRLDLDAKNNLETYSRERSRQAESTTD 453
Db 767 SLDVKISEAMNNKFTRECSVTY-----LFKNWLPKVIDELNK--FDLTKTELINLID 819
Qy 454 PMEMTNGAGGEGYSSNSITVAQVAYDNLRI-----NNILEDL 496
Db 820 ------SHNIIUVGEV-----DLKAKVNESPENTMPPNIFSYTNNSLLKDI 860

Qy 497 -----SKAWCREQHRALV-----WNLSKINPTSVMSMIYNRPVSAKRIGD--VI 540
Db 861 INEYFNSINDSKLSLQKKNALVDTSGYNAEVRVGNVOLNTIYTNDKFLUSSGDKIIV 920
Qy 541 SVSNCI-----VVDOTSVSL-----HKSRLLSASDEK-----CFSRPPVTFKFM 580
Db 921 NLNNILYSAIYENSSVSFWIKSKDUTNSHNEVTIINSIQNSGWLCTIRNGNIWILQ 980
Qy 581 NDSTIYKQGLGVANNEILLTTTLETQCENTEYFQAKTD-----MYTKNYEHLKTVPLS 635
Db 981 DVNRKYSLIFDYSESLSHTGY-----TNKWFVTTNNIMGYMKLYINGELKQSKIE 1034
Qy 636 SITLDTFIALNFTLLENVDFKVIELYTRDEKLSNVFDIETMFEYN--YYAQRVSGLR 693
Db 1035 DLDEVKLDKTIIVFGIDENIQM-LWIRD-----FNIFSKELSGNEINIVYEGQILRNVI 1089
Qy 694 KD 695
Db 1090 KD 1091

RESULT 23

PCT-US02-24459-102
; Sequence 102, Application PC/TUS0224459
; GENERAL INFORMATION:
; APPLICANT: Curagen Corp. et al.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-416C-061 (Cura-716 SMT WO)
; CURRENT APPLICATION NUMBER: PCT/US02/24459
; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: 60/309,501
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 60/316,508
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: 60/354,655
; PRIOR FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: 60/310,291
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 60/383,887
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: 60/310,951
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/323,936
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: 60/381,039
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/311,292
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/311,979
; PRIOR FILING DATE: 2001-08-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 102
; LENGTH: 1664
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-24459-102

Query Match 2.7%; Score 122; DB 1; Length 1664;
Best Local Similarity 18.0%; Pred. No. 27;
Matches 133; Conservative 97; Mismatches 201; Indels 306; Gaps 34;
Qy 151 DSITNOHTFYKSIETPETERKMDTIYQCFNS-LRLNTGGNLLTYVDRDDINMTVFLQPDVG 209
Db 191 DSVHQLHVYHRDIKP-----DNILMDMNGHIRLADFGCLKLMDEGTQVSSVAVGTFDY 244
Qy 210 VTPDV-----KRYGSOPE-----LYLEPGWF-----WG---SYRRR--- 237
Db 245 ISPEILQAMEDGKRGYRGPEDCMWMLGVMCYMELYGETPFYAESLVETIYKIMNHRKRFQF 304
Qy 238 ----TTVNCLEMDMPAR-----SNPPDFPVATGDTVMSFPWSGEDDHENKMH 283

Db 305 PAQVTDVSENADLIRRLTCSREHRLGQNGIEDF : ||: ||: :
Qy 284 EKPWFVSV-----INNVKVD--YONRGTVP-----KKHPFFSGIDWDNIRNC 355
Db 356 EAPYIPEVSSPTDTSNFDVDDCLKNSETMPPHTAFSGHLPFVGTYTSSCVLSDRS 307
Qy 308 -----LGKTRI-----FLDREYTLSEKHLKMSYCPITLWKAFYNGIQTEHSGSY 354
Db 416 CLRVTAGPTSLDLVNQVORTLDNNLATEAYERRIKRLEQEKLELSRKQESTQTQVALQY 475
Qy 355 HFVANDITASFTSKEDMKFNTYHCLNEBK-----AEIEKKYAKVNSTHSGY 405
Db 476 STVDGLPTAS-----KDEIKN-----LKEBIEKLRKQVTESSHLEQOOLEANAVRQELD 525
Qy 406 DL-----KYFKTDGGLVLMQPLQNR-----LLDAKNKLN----- 437
Db 526 DAFROIKAYEKQIKT-----LOQERDLNKLQVQSERLKNQSKELKDAHCQRKL 575
Qy 438 -----ETYSRRSR-----ROAESTTDPMM-----BMT 459
Db 576 AMQEFMEINERLTHTQKQKLARHVRDKKEEVDLMQKVESLRQELRRTERRAKKELEVH 635
Qy 460 GNGAGGEYSSENSITVAQVQAYDNLRINNILEDL-----SKAWCRQOHRALVWN 512
Db 636 TEALAAEASKDRKLRQESHEYSK-----QLENELEGLKQKQISYSPGVCSEHQ-----Q 685
Qy 513 ELSKINPTSVMSMIYNRPVSARKRIGDVISVSNICVVDQTSVSLH-----KSLRLLSASDEK 568
Db 686 EITKLTDLKKSIFYEEELSKREG-----IHANEIKNLK----- 720
Qy 569 CFSRPPVTFKFMNDSTIYKQ-QLGVNNEIL-----LTTTYLETQENTYFYFOAKTMYI 622
Db 721 -----KELHDS-----EGQALNKEIMILKDKLEKTRRESQSEREFSEFK----- 763
Qy 623 YKNEYHLKTVPLUSSITLDTFIALNFTLLENVDFKVIELTRDEKRLSNVFD-IETMPRE 681
Db 764 -QOYEREKV-----LTTTEENKLTSELDKLTTLTYEN 793
Qy 682 YNYVAQRVSGLRKDLDD 698
Db 794 LSIHQOLEEVEVKDLAD 810

RESULT 24
PCT-US02-24459-100
; Sequence 100, Application PC/TUS0224459
; GENERAL INFORMATION:
; APPLICANT: CuraGen Corp. et al.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-416C-061 (Cura-716 SMT WO)
; CURRENT APPLICATION NUMBER: PCT/US02/24459
; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: 60/309,501
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 60/316,508
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: 60/354,655
; PRIOR FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: 60/310,291
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 60/383,887
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: 60/310,951
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/323,936
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: 60/381,039
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/311,292
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/311,979
; PRIOR FILING DATE: 2001-08-13

; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 100
; LENGTH: 1738
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-24459-100

Query Match 2.7%; Score 122; DB 1; Length 1738;
Best Local Similarity 18.0%; Pred. No. 28;
Matches 133; Conservative 97; Mismatches 201; Indels 306; Gaps 34;

Qy 151 DSITNQITFYKSIPIWETEKMDTIYQCFSN-LRLNTGNNLLTYVDRDINNMTVFLQPVVDG 209
Db 191 DSVHQLHVHRDIKP-----DNILMDMNGHRLADFGSCLKLMEDGTQVQSSVAVGTPTY 244
Qy 210 VTPDV-----KRYGQPE-----LYLEPGWF-----WG-----SYRRR--- 237
Db 245 ISPEILOAMEDGKGRYGPECDSWLSGLVCMYEMLYGETPFYAESLVETVYKIMNHKRFQF 304
Qy 238 -----TTNCELMDMFAR-----SNPPDFEFTVATGDTVEMSPFMSGCDHDKMH 283
Db 305 PAQVTDVSENADLIRRLTCSREHRLGQNGIEDF-----KHPFFSGIDWDNIRNC 355
Qy 284 EKPWFVSV-----INNVKVD--YONRGTVP----- 307
Db 356 EAPYIPEVSSPTDTSNFDVDDCLKNSETMPPHTAFSGHLPFVGTYTSSCVLSDRS 415
Qy 308 -----LGKTRI-----FLDREYTLSEKHLKMSYCPITLWKAFYNGIQTEHSGSY 354
Db 416 CLRVTAGPTSLDLVNQVORTLDNNLATEAYERRIKRLEQEKLELSRKQESTQTQVALQY 475
Qy 355 HFVANDITASFTSKEDMKFNTYHCLNEBK-----AEIEKKYAKVNSTHSGY 405
Db 476 STVDGLPTAS-----KDEIKN-----LKEBIEKLRKQVTESSHLEQOOLEANAVRQELD 525
Qy 406 DL-----KYFKTDGGLVLMQPLQNR-----LLDAKNKLN----- 437
Db 526 DAFROIKAYEKQIKT-----LOQERDLNKLQVQSERLKNQSKELKDAHCQRKL 575
Qy 438 -----ETYSRRSR-----ROAESTTDPMM-----BMT 459
Db 576 AMQEFMEINERLTHTQKQKLARHVRDKKEEVDLMQKVESLRQELRRTERRAKKELEVH 635
Qy 460 GNGAGGEYSSENSITVAQVQAYDNLRINNILEDL-----SKAWCRQOHRALVWN 512
Db 636 TEALAAEASKDRKLRQESHEYSK-----QLENELEGLKQKQISYSPGVCSEHQ-----Q 685
Qy 513 ELSKINPTSVMSMIYNRPVSARKRIGDVISVSNICVVDQTSVSLH-----KSLRLLSASDEK 568
Db 686 EITKLTDLKKSIFYEEELSKREG-----IHANEIKNLK----- 720
Qy 569 CFSRPPVTFKFMNDSTIYKQ-QLGVNNEIL-----LTTTYLETQENTYFYFOAKTMYI 622
Db 721 -----KELHDS-----EGQALNKEIMILKDKLEKTRRESQSEREFSEFK----- 763
Qy 623 YKNEYHLKTVPLUSSITLDTFIALNFTLLENVDFKVIELTRDEKRLSNVFD-IETMPRE 681
Db 764 -QOYEREKV-----LTTTEENKLTSELDKLTTLTYEN 793
Qy 682 YNYVAQRVSGLRKDLDD 698
Db 794 LSIHQOLEEVEVKDLAD 810

RESULT 25
US-10-679-063-11928
; Sequence 11928, Application US/10679063
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52054)B

```
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52054)B
; CURRENT APPLICATION NUMBER: US/10/679,063
; PRIOR FILING DATE: 2003-10-02
; PRIOR APPLICATION NUMBER: 60/415,758
; NUMBER OF SEQ ID NOS: 27373
; SEQ ID NO 11928
; LENGTH: 836
; TYPE: PRT
; ORGANISM: Plasmodium falciparum 3D7
US-10-679-063-11928

Query Match      2.7%; Score 121.5; DB 6; Length 836;
Best Local Similarity 16.8%; Pred. No. 15;
Matches 129; Conservative 137; Mismatches 255; Indels 245; Gaps 39;

Qy 10 SVLAWLYOVALYSLIAETGVTSPNTA-TWSTESPLTGHYGTHTDSSHGERGNENRD 68
Db NVLCL-ELFFLSLFFIMI-----ISADNNTAHTFTSSPTNNNNNNNNNTDNDNSNNNS 90
Qy 69 SEEQNK-----NIYGSPTFPYRVCS---ASGVGVF--RFQTDHVCPPASP 110
Db 91 NNNKNNYINNHHNNHNNYFDDPSIKYRKNNNIYKDNINDEFLLKKEDNDIMNNMLD 150
Qy 111 MV-----HSGILLIYKQ---NIIPFMFRVKYRKVVTTSTVNGI-----YSD 151
Db 151 NILIOCLHFKGAYKIRKNNKNNMSDGIPEA---KLVTKLALNNVTSYIFFYFNPYDN 206
Qy 152 SITNQHTFYKSIPEWTEKMDTIYQCFNSRLNTGGNLLTYVDORDDINMTVFLQPVGVGT 211
Db 207 NNTNHNIIYILNDIPKQYK---YLGFGNINKENRYI--FNGQGYNISTF----- 251
Qy 212 PDVKRYGSOPELYLEPGWFSYRRRTTVNCELMDMFARSNPDPDFVTATGDTVMSPP 271
Db 252 -----COLKNNVKKNEAYKYAHT-----NI 272
Qy 272 WSGDDHDK-----MHEK---PWFVSV---INNYKVVDD---YQNRGTVPGLK 310
Db 273 YNNNDNNNNNNINGVNECLNYTINQYINKYESVLSREVDESYSLSLKYI----- 328
Qy 311 TRIFLDREYVTLSEKHLKMSYCPILWKAFYNGIQTEHSGSYHFVANDITASFTTSKE 370
Db 329 ----LDEQY---YKNYLQ--EY--IKLYKNYISYIQYVHTHY-----IEKE 366
Qy 371 DMKEFNITYHCLNEIEIKAEIEKKYAKVNSTHYSKYGDLKYFKTDGGLYLVNQPLIQNRLD 430
Db 367 NMIFHSLGNVNEEIIKHQKLSAKQNDKXLSVYTLKKVKBESG---TIPNHT 419
Qy 431 AK-----NKLNNETYSRSRROAESTTDPMMEMTNGAGGEYSSENSITVAQVQYAYD 483
Db 420 REYTSNVNDNGNNKNDGNNKNDNNNSDGNNDGNNKNDYNN-----KYN 471
Qy 484 NLRIRINILEDLSKACWCRQRAALVNLNELSKINPTSVMSIYNRPVSARIGDVISVS 543
Db 472 D-----VNNKNDVNNKY--NDGH-----NKNNEHVTNAHYD-----GYIYS-N 508
Qy 544 NC-IWVDQTSVSLHKSRLLSASDEKCFSPPTVF-----KPNNDSTIYKQGVNNEI 596
Db 509 NCNLLISFEGIDI-----DKYISTVKLVNFSILFNKISIEISLFSWQIGTSGSI 558
Qy 597 --LITTYTLETQCENTEYVFOAKTDMVIYKNEYHLKTVPLSSITTLDTFIA-----LNFT 649
Db 559 PGASRISLISICLSNLDIFESULLLY-----EVLSSKLLIIFILMILLKFLFLT 609
Qy 650 LLE-----NVDPKVIETYTRDEKRLSNVFDIETMFREYNY 685
Db 610 IMEVRYVLIWKAHQHVEGWEHQMOKLSKL-----YKYY 646

RESULT 26
US-10-679-063-18214
; Sequence 18214, Application US/10679063
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
```


QY 819 QONSHMETKTRKDPKDSIL 837
Db 1086 --NNHINLNTKKKKGNVL 1102

RESULT 27
PCT-US03-27401-364
; Sequence 364, Application PC/TUS0327401
; GENERAL INFORMATION:
; APPLICANT: TUFTS UNIVERSITY
; TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE FOR DIAGNOSIS, TREATMENT AND
; PREVENTION OF ACTIVE INFECTION
; FILE REFERENCE: 700355-52941-PCT
; CURRENT APPLICATION NUMBER: PCT/US03/27401
; PRIOR FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US 60/407,082
; PRIOR FILING DATE: 2002-08-30
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 364
; LENGTH: 2004
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
PCT-US03-27401-364

Query Match 2.7%; Score 121; DB 1; Length 2004;
Best Local Similarity 18.5%; Pred. No. 36;
Matches 165; Conservative 108; Mismatches 295; Indels 322; Gaps 43;

QY 28 AETGVTSPNTATWSTESPLTGHYTHDSHGGERGNENRDEE--QNKNIYGSPTFPY 85
Db 669 SENGQTEPEPSNGNSTED-----VSTESNTSNGNEEIKQENELDPDKKVEEPEKTLLEL 723

QY 86 RVCSASGVDFRFTQDHVCPDASDMVHSEGLILLYKQNI-----IP-----FMFRVRKYR 136
Db 724 R-----NVSDLELYSL-----SNG---TYKHISLEQVPSNPNSFYFVKV--- 760

QY 137 KVVTTSTVYNGIY-----SDSITNQHTFYKSIPEWTEKMDTIYQCFNSLRNLNTGNLL 190
Db 761 -----SSSFQDVLPVASISEERKNDKILYKI-----TAKVEKLOQ---ETESRYKDNFT 807

QY 191 TYVDR-----DDINMTVFLQPDVDPVKRYGSOPELYLEPQFWGFSYRRRTTVNCELMD 246
Db 808 FYLAKKGTEETNTFTSF-----SNLVKAINQNPS-----GTVHLAASLANANEVE 851

QY 247 M-----FARSNPPDFEFTVATGDTVE-----MSPFWSGEDD- 277
Db 852 LGPDRSYIKDTFTGRLIGEKDGKNYAIYNLKKPLFENLSGATVEKLSKNVAISGKDDI 911

QY 278 -----HKNMKHEKPFVSVINNYKVVDYQN 302
Db 912 GSLANEQNNTKIKQVHVDGVLAGEGIGGLLAKAEQSSITESSFKGRIINTYETTAYN 971

QY 303 RG-----GKTRIFLDRE-----EYTLISWE 325
Db 972 ICGMVGHLTGDKALLTKSKATVAISSNTNTSDQTVGGLAGLVDRDAQIDSYAEGDINN 1031

QY 326 KHLKMSYCPLTLWKAIFYNGIQTSHSGS-----YHFVANDIATSTT 367
Db 1032 RHFGRVAGVAGNLWDRDTSQDVR--HAGSLTNVLSDVNVNTNGNAITGYHY--NEMKVKDTF 1087

QY 368 SKEDMKFEFTTTHCLNEEI-----KAEIEKKYAKVN-----STHSK 403
Db 1088 SSKANRVYNT--LVKDEVWSKESFEERCTMLDASQIAKAEINPLILPTVEPLSTSGK 1145

QY 404 -----YDGLKYFKTDGGL-YLVNQPLI-----QNRLLDAKNKL----- 435
Db 1146 KDSDFSKVAYYQAKNLTYNKIEKLLPFYNKATIVKYGNLVNNSLLYQKELLSAVMMKD 1205

QY 436 -----NNYTSRRSRQAESTTDPMMETGNCAG--GEYSSENS-ITVAQVQVAYD 483
Db 1206 NOVIITDIVSNKQATANKLLHLYKDDLSSEKLDLKYQNDFAKLAEYLSGNTGLLTPNQFLYD 1265

QY 484 NLRIRINNILEDLSKAWCREOHRALVWNLKINPTSVMSIYNRPVSAKRIGDVISVS 543
Db 1266 QTSI-IKQVLPDLQKV---DYHSEAI--RKTGLGISPNVKQTELYLEQDFAK----- 1310

QY 544 NCIVVDOTSUSLHKSRLLSASDEKCFSPRPVTFKFNNDSTIYKGOGLGVNNE-ILLITTY 602
Db 1311 -----TKQLEDLSLKKLLSADAGLASANPVTGELYVD-----KIKRKEALLGLUTY 1357

QY 603 LETQCENTEYFQAKTDMYIYKNYEHKTVPLSSITITLDTFIAL-----NFTLLENVDF 656
Db 1358 LERMY-NFSYQVNVKDLVLY---HLDFFGCKGNASPLDTLIELGKSGFNLLAKNNVDT 1412

QY 657 KVIELYTRDEKRLNVSFDIETMPREYNYAORVSGRLKOLLDLSTNRNQF 706
Db 1413 YGIS-----LASQHTTDLFTLEHY-----RKVFLPNTSNDWF 1447

RESULT 28
US-10-472-928-2306
; Sequence 2306, Application US/10472928
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; TITLE OF INVENTION: THE INSTITUTE FOR GENOMIC RESEARCH
; TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026926WO
; CURRENT APPLICATION NUMBER: US/10/472,928
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: GB-0107658.7
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 4979
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 2306
; LENGTH: 2004
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; OTHER INFORMATION: Immunoglobulin A1 protease (iga)
; OTHER INFORMATION: Cellular location: Peptidoglycan-bound (LPXTG)
; OTHER INFORMATION: Similar to strain R6 sequence 15903086 (O.E+01)
US-10-472-928-2306

Query Match 2.7%; Score 121; DB 6; Length 2004;
Best Local Similarity 18.5%; Pred. No. 36;
Matches 165; Conservative 108; Mismatches 295; Indels 322; Gaps 43;

QY 28 AETGVTSPNTATWSTESPLTGHYTHDSHGGERGNENRDEE--QNKNIYGSPTFPY 85
Db 669 SENGQTEPEPSNGNSTED-----VSTESNTSNGNEEIKQENELDPDKKVEEPEKTLLEL 723

QY 86 RVCSASGVDFRFTQDHVCPDASDMVHSEGLILLYKQNI-----IP-----FMFRVRKYR 136
Db 724 R-----NVSDLELYSL-----SNG---TYKHISLEQVPSNPNSFYFVKV--- 760

QY 137 KVVTTSTVYNGIY-----SDSITNQHTFYKSIPEWTEKMDTIYQCFNSLRNLNTGNLL 190
Db 761 -----SSSFQDVLPVASISEERKNDKILYKI-----TAKVEKLOQ---ETESRYKDNFT 807

QY 191 TYVDR-----DDINMTVFLQPDVDPVKRYGSOPELYLEPQFWGFSYRRRTTVNCELMD 246
Db 808 FYLAKKGTEETNTFTSF-----SNLVKAINQNPS-----GTVHLAASLANANEVE 851

QY 247 M-----FARSNPPDFEFTVATGDTVE-----MSPFWSGEDD- 277
Db 852 LGPDRSYIKDTFTGRLIGEKDGKNYAIYNLKKPLFENLSGATVEKLSKNVAISGKDDI 911

QY 278 -----HKNMKHEKPFVSVINNYKVVDYQN 302
Db 912 GSLANEQNNTKIKQVHVDGVLAGEGIGGLLAKAEQSSITESSFKGRIINTYETTAYN 971

QY 303 RG-----GKTRIFLDRE-----EYTLISWE 325
Db 972 ICGMVGHLTGDKALLTKSKATVAISSNTNTSDQTVGGLAGLVDRDAQIDSYAEGDINN 1031

```
QY 326 KHLKMSYCLTLWKAFYNGIQTEHSGS-----YHFVANDITASFTT 367
|||
Db 1032 KHFRVAGVAGNLDRTSGDVR--HAGSLTNVLSVDVNTNGNALTGVHY--NEMKVKDTF 1087
|||
QY 368 SKEDMKEFNTYHCLNEEI-----KABIEKKYAKVN-----STHSK 403
|||
Db 1088 SSKANRVYNT--LVKDEVVSKESFEERGTMLDASQIASKABINPLILPTVEPLSTSGK 1145
|||
QY 404 ----YGLDKYFKTDGGL-VLWQPLI-----ONRLDANKL-----435
|||
Db 1146 KDSDFSKVAYYOAKRNLTYNKIEKLPPFYNKATIVKYGNLVNNSLSYQKELLSAVMMKD 1205
|||
QY 436 -----NNEYTSRRSRQAESTTDPMMEMTCGAG--GEYSSENS-ITVAQVQVAYD 483
|||
Db 1206 NQVITDIVSNKQANKLLHYKDDLSEKLDLKQNDPAKLAESLGNLTGTYPNQFLYD 1265
|||
QY 484 NLIRINNILEDISKWCROHRAALVWNLKINSKINPTSVMSIMYRNPVSAKRIGDVISVS 543
|||
Db 1266 QTSI-IKQVLPDLQV--DYHSEAI--RKTGIGSPNVKQTELYLEDOFAK-----1310
|||
QY 544 NCIVDQTSVSLHKSLELLSASDEKCFSPRPVTFKPMNDSTIYKQGLGVNNE-ILLTTY 602
|||
Db 1311 -----TKOOLEDSLKKJLSADAGLASANPVTGYLVD-----KIKRKEALLGLTY 1357
|||
QY 603 LETCOENTYFFOAKTDMYIKNYEHLKTVPLSSITLDTFIAL-----NFTLLENVDF 656
|||
Db 1358 LERWY-NFSQOVNVKDLVY----HLDFFGKGNASPLDTLIELGKSGFNNLAKNVDT 1412
|||
QY 657 KVIELYTRDEKRLSNVFDIETMFREYNYAQRVSGLRKDLLDSTNRNQF 706
|||
Db 1413 YGIS-----LASQHTTDLFTLEHY-----RKVELPNTSNNDF 1447
|||

RESULT 29
US-10-679-063-18007
; Sequence 18007, Application US/10679063
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52054)B
; CURRENT APPLICATION NUMBER: US/10/679,063
; CURRENT FILING DATE: 2003-10-02
; PRIOR APPLICATION NUMBER: 60/415,758
; PRIOR FILING DATE: 2002-10-02
; NUMBER OF SEQ ID NOS: 27373
; SEQ ID NO 18007
; LENGTH: 846
; TYPE: PRT
; ORGANISM: Plasmodium yoelii yoelii
US-10-679-063-18007

Query Match 2.7%; Score 120.5; DB 6; Length 846;
Best Local Similarity 16.8%; Pred. No. 17;
Matches 135; Conservative 114; Mismatches 265; Indels 289; Gaps 34;

QY 16 YLYOVAL-YSLSIAETGVTSPNATWSTES-PLTGHY--GTHDSSHGGERGNENRDEE 71
|||
Db 134 YLTQIGLVYLLSKENSIMCEYAKVQNNMLPCLNWINSSHNSYLGRK-----183
|||
QY 72 QNKNIYSPSTFPYRVCSAGSGVDVFRFQTDHVCPCDASDMVHSEGIILLYKQNIIPFMR 131
|||
Db 184 -----QIFSAISIE-----QYIYILDCRCVEFD--CYVFKNIV-----217
|||
QY 132 VRKRVKVVTSVYNGIYSSITNQHTFYKYSIEPWETERKMDTIYQC-----FN 179
|||
Db 218 -----VYHGLYGYKLTSSILF-----CDTLACMFGFSTSPYIIL 254
|||
QY 180 SLRL-----NTGGNLLTVVDRDDINMTVLPQVDGVTDPVKRYGSOPELYLEPGFWG 232
|||
Db 255 SLEIHCCKKHLNLIKLIYILGNQLYIPKTKDEINNIPTN-----295
|||
QY 233 SYRRRTTVNCELMDMFARSPPDFPFVATGDTVEMSPFWSGEDDHENKMH-EKPMFVSV 291
|||
```

```
Db 296 -----NCR--NKFVLVKYKHFDNNYT-----SSFYSIFEGUKVMVYDELGYFSOI 337
QY 292 IN-----NYKVDYQNRGTVPGLGKTRIFLDREEEYTLSEWEHLKN 330
|||
Db 338 INEQEDIDVDEPKEYOIIIEKKNEKYNKISIEQNOMPEDEPVNNYIE-----QNNSKN 389
|||
QY 331 MSYCPLTLWKAFYNGIQTEHSGSYHFVANDITASFTTSKEDMKFNTTY-----HCLNE 384
|||
Db 390 TSYKNNQVQKTNKNSDQ-----NNLCNGENVSQDRHEVDNSYSSSSSSNCLYE 438
|||
QY 385 EIKABIEKKYAKVNSTHSHKYGDLYKPYKTDGGLYLVWQPLIONRLDANKLNNETYSRRS 444
|||
Db 439 N-----NTIHKTIQIDEKYRND-----NCSN--NNSGFNNS 468
|||
QY 445 RROAESTTDPMMEMTCGAGGEYSSENSITVAQVAYDNLRIRINNILEDLS--KAWCR 502
|||
Db 469 VPNTNTSOSI-----TNCCKMSETAHMSKEBYKKQDNEKYK--NNILKNCCLKGYVF 522
|||
QY 503 EQHRAALVWNLKINSKINPTSVMSMI-----YNRPV-----SAKRIGD-----538
|||
Db 523 RNFQENRTYNEICISISENKFILIKKNEDDVIKYNQKTLTRVYPSGTRLASTNPNLIFW 582
|||
QY 539 ----VISVNCIVVDQTSVSLHKSRLLSASDEKCFSPRPVTFKPMNDSTIYKQGLGVNNE 595
|||
Db 583 SAGIQFVALNYQVNGLSMLLNKGRFLENGGKHSGYILKPELLRF-NKKHDY-NVLSLNIQ 640
|||
QY 596 IL-----LTTY-----LETCOENTYFFOAKTDMYIKNYEHLKT 631
|||
Db 641 ILSLHQLNLLFSIKNYQEKKLKELFKMDMIQRIOTHKKISNKRKKFKYLOKLREKKS 700
|||
QY 632 VPLSSITLDTFIALNFTLLENVDFKVIELYTRDEKRLSNVFDIETMFREY-----NY 684
|||
Db 701 PLLSDIQSD-----SKKKKISN-----ETLLKKFKDNDNDVNY 735
|||

QY 685 YAQRVSGLRKDLLDSTNRNQFV 707
|||
Db 736 IHKKFANIEKEYEDMLTEYKSFL 758
|||

RESULT 30
US-10-679-063-18197
; Sequence 18197, Application US/10679063
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52054)B
; CURRENT APPLICATION NUMBER: US/10/679,063
; CURRENT FILING DATE: 2003-10-02
; PRIOR APPLICATION NUMBER: 60/415,758
; PRIOR FILING DATE: 2002-10-02
; NUMBER OF SEQ ID NOS: 27373
; SEQ ID NO 18197
; LENGTH: 793
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-10-679-063-18197

Query Match 2.6%; Score 120; DB 6; Length 793;
Best Local Similarity 18.1%; Pred. No. 17;
Matches 106; Conservative 90; Mismatches 237; Indels 152; Gaps 25;

QY 315 LDREYTLGWEKHLKMSYCLTLWK-----AFYNGIQTEHSGSYHFVANDITASFT 366
|||
Db 32 LARKNNMTWGGYGGYEESEITICRENGLVESFQFVENVENPH-----VNDTASLSFL 85
|||
QY 367 TSKEDM--KEFNTTYHCLNEEIKABIEKKYAKVNSTHSHKYGDLYKPYKTDGGLYLVWQPLI 424
|||
Db 86 SCSCNCIVTRMLNHHNVIVEDIILNEHKDKYANVLFNFPVYNNNE--KVKNGIYEDY--LL 141
|||
QY 425 QNRLDAA-----KNKLNNETYSRRSRQAESTTDPMME 457
|||
Db 142 NSRLLVTVNSTGHVHLNWEHDESSKYVNDTLTKNEIKNEDIQKKMQRYCENCYKQVNLN 201
|||
```



```
Db 236 FSYEPQAGTVRVVNYRTDGNVYKSLWYMGDKVKNPSS-----AQWPDGTF--TA 284
Qy 262 TG-----DTVEMSPFWSGEDDHENKHEKMPFVSVNNYKVDYQNRG 304.
Db 285 TGKGRYIDIPLENAAREFGFLLDDESKQGGDDVKIRK-----ENYKFTDLKNH- 332
Qy 305 TVPLGKTRIFL---DREYTLSEWEKHLKNMSYCLTLWKAFYNGIQTEHSGSVHFVANDI 361
Db 333 -----SQIFLKDDDESITYNPVYVHDIRM-----TGAQHVGTSSI 367
Qy 362 TASFTT---SKED-MKEFNTTYHCLNEEIKAEIEKKYAKVNSTHS--KYGDLYK- 409
Db 368 ESSFSTLVGAKKEDILKHSNITNLGNKVITTDVAIDEAGKVKVYSGDFDKHPYTVSY 427
Qy 410 -----FKTDGGLYVWQPLIQNRLLDNAKKNLNNETYS-----RRSRQAEST 451
Db 428 NSDQFTTKTS---W-----RLKDETSYDGLKADLKEEGKQVDLTLSWSPS 470
Qy 452 -----TDP-----MMEMTGNAGGE--YSSSENSITVAQVQYAYDNLRIINN- 491
Db 471 ADKVVVVYDKNDPKVGVTVALEKBERGTWKQTLDSNKLGITDFTGYGYQYQIERQKG 530
Qy 492 ---ILEDLSK---AWCREQHR-----AALVMNELSKINPTSVMSMIYNRPVSAKRIGD 538
Db 531 TVLALDPYAKSLAAMNSDDSKIDDAHKVAKAAFPDPAKLGP---QDLTYGKIHNFKTRD 587
Qy 539 VISVNCIVVDQTSVLSHLKSLRLLSASDEKCFSPRPVTFKPMNDSTIYKQGLGVNNEILL 598
Db 588 AV-IYEAHVDRFTS-----DPAIAKDLTKPFGTFEAFIEKLDYLDKDLGVTHIQLL 636
Qy 599 TTYLETQCENTEYFFOAKTDMYIYKNYEHKLTVPPLSSITTLDTFIALNFTLLENVDVKV 658
Db 637 PV-----LSYFVNE-----LKNHERLSDYASSNSNWNWGYDPQNY-----FSL 675
Qy 701 TNRNQFVDAFGSLMDDLG 718
Db 727 PNYHFMDADGTPRTSFG 744
```

RESULT 33

```
US-10-472-928-364
; Sequence 364, Application US/10472928
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: THE INSTITUTE FOR GENOMIC RESEARCH
; TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026926W0
; CURRENT APPLICATION NUMBER: US/10/472.928
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: GB-0107658.7
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 4979
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 364
; LENGTH: 1280
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; OTHER INFORMATION: alkaline amylopullulanase, putative
; OTHER INFORMATION: Cellular location: Peptidoglycan-bound (LPXTG)
; OTHER INFORMATION: Similar to strain R6 sequence 15902291 (0.E+01)
```

```
Query Match 2.6%; Score 118.5; DB 6; Length 1280;
Best Local Similarity 20.3%; Pred. No. 32;
Matches 174; Conservative 102; Mismatches 297; Indels 285; Gaps 47;
Qy 9 QSVL--ALWLYQVALYSLSLAETGVTS--PPNTATWSTESPLTGHYGTGSHSGRGNN 64
```

```
Db 24 GSVILGASLVLLAMATPTTISSDESTPTTNEPNNRNTTTLAQPLT-----DTAAGS-GKN 76
Qy 65 ENRDSEQKN-----IYGSPSTFPYRVCSASGVDFRQTDHVCPSDASDMVHSEGILL 119
Db 77 ESDISSFGNANASLEKTEEKPAASPADPAPQTG-QDRSSEPTTSTSPVTTETKAEPIED 135
Qy 120 IYKQNIIPMERVRKRVKVVTTSTVNGIYS-DSITNQHTFYKSIIEPW-----ETEK 170
Db 136 NY-----FRIH-VKLUPEENKDAQGLWTWDVE-----KPSNWNPGALSFKDAKK 180
Qy 171 MDTIYQCFNSLRL-----NTGNNLIT-----YVDRDDINMTV 202
Db 181 DD--YGYLDVCLKGFOAKKISFLINNTAGKNLTGDKSVEKLVPKMNEAWLQDQ---YKV 235
Qy 203 F-LQPVGDVTPDKRVYGSQPELYLEPCWFGSVYRRRTTVNCELMDFAKSNPFPDEFVTA 261
Db 236 FSYEPQAGTVRVVNYRTDGNVYKSLWYMGDKVKNPSS-----AQWPDGTF--TA 284
Qy 262 TG-----DTVEMSPFWSGEDDHENKHEKMPFVSVNNYKVDYQNRG 304
Db 285 TGKGRYIDIPLENAAREFGFLLDDESKQGGDDVKIRK-----ENYKFTDLKNH- 332
Qy 305 TVPLGKTRIFL---DREYTLSEWEKHLKNMSYCLTLWKAFYNGIQTEHSGSVHFVANDI 361
Db 333 -----SQIFLKDDDESITYNPVYVHDIRM-----TGAQHVGTSSI 367
Qy 362 TASFTT---SKED-MKEFNTTYHCLNEEIKAEIEKKYAKVNSTHS--KYGDLYK- 409
Db 368 ESSFSTLVGAKKEDILKHSNITNLGNKVITTDVAIDEAGKVKVYSGDFDKHPYTVSY 427
Qy 410 -----FKTDGGLYVWQPLIQNRLLDNAKKNLNNETYS-----RRSRQAEST 451
Db 428 NSDQFTTKTS---W-----RLKDETSYDGLKADLKEEGKQVDLTLSWSPS 470
Qy 452 -----TDP-----MMEMTGNAGGE--YSSSENSITVAQVQYAYDNLRIINN- 491
Db 471 ADKVVVVYDKNDPKVGVTVALEKBERGTWKQTLDSNKLGITDFTGYGYQYQIERQKG 530
Qy 492 ---ILEDLSK---AWCREQHR-----AALVMNELSKINPTSVMSMIYNRPVSAKRIGD 538
Db 531 TVLALDPYAKSLAAMNSDDSKIDDAHKVAKAAFPDPAKLGP---QDLTYGKIHNFKTRD 587
Qy 539 VISVNCIVVDQTSVLSHLKSLRLLSASDEKCFSPRPVTFKPMNDSTIYKQGLGVNNEILL 598
Db 588 AV-IYEAHVDRFTS-----DPAIAKDLTKPFGTFEAFIEKLDYLDKDLGVTHIQLL 636
Qy 599 TTYLETQCENTEYFFOAKTDMYIYKNYEHKLTVPPLSSITTLDTFIALNFTLLENVDVKV 658
Db 637 PV-----LSYFVNE-----LKNHERLSDYASSNSNWNWGYDPQNY-----FSL 675
Qy 659 IELYTRD---EKRLSNVFDIETMFE-----YNYAQRVSGLRKOLL-DLS 700
Db 676 TGMYSDDPKNPEKRIA---EFKNLINEIHKMGAILDVVYNHTA-----KVDLPEDLE 726
Qy 701 TNRNQFVDAFGSLMDDLG 718
Db 727 PNYHFMDADGTPRTSFG 744
```

RESULT 34

```
PCT-US03-30720-1209
; Sequence 1209, Application PC/TUS0330720
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Zhou, Ping
; APPLICANT: Zhang, Jie
; APPLICANT: Ghosh, Malabika
; APPLICANT: Ren, Feiyan
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhao, Qing A.
```


QY 736 SSIV--TGFINFKNPGGMLIIWVGVLFAIVFLTKTKIYETAPIKMIYPEIDKLKE 793
Db 715 DKVETESKLOSHNN-----THEAQMSIEDFQWYEEQSKSEE 754
QY 794 REGKSEIAP-----SEELERIVLA-----MHIHQNSHME--TKTRKOPKDS 835
Db 755 QKLLADISSLVSKHITRQRELGVRLSSGLDGAARGKAFDEHTSAMFVTDKAKRKWET 814
QY 836 ILTRAQNMILKRSGYNLKNKNAESVEML 862
Db 815 FAEQAENDCKAGSSSAVKHCRMETML 841

RESULT 36
US-10-330-773-956
; Sequence 956, Application US/10330773
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001300
; CURRENT APPLICATION NUMBER: US/10/330,773
; CURRENT FILING DATE: 2002-12-27
; NUMBER OF SEQ ID NOS: 981
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 956
; LENGTH: 880
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-330-773-956

Query Match 2.6%; Score 117; DB 6; Length 880;
Best Local Similarity 17.5%; Pred. No. 28;
Matches 113; Conservative 105; Mismatches 239; Indels 190; Gaps 24;

QY 213 DVKRYGSOPELYLEPGWFGWSYRRRTTNCNCELMDFARSNPPDFVTATGDTVEMSPFW 272
Db 362 EAEHHSSPKRELPFG-----LGDNAEL---M 385
QY 273 SGEDDHENKMKHPWFVSVIN-NYKVVDYQNRGTVPGLGKTRIFLDREYTLSEKHLKNN 331
Db 386 GGQDQMDERDVRFRQLKAELNSVIRKLEDRNTLLADERNELKRSRET---EVOLK-- 440
QY 332 SYCPLTLWKAFYNGIQTEHSGSVHFVANDITASFTTSKEDMKEFNTTYHCLNEIEKAE 391
Db 441 ---PL-----VEKNRMKNKNEELLHSIQRMEEKLSLTRENVMEMKEKLSAQAS 486
QY 392 KKYAKVNSTHSGYGLKYFTDGLG-LYVMOPLIQNRLDLAKNKLNNETYSRRSRQAES 450
Db 487 LK-----RHTSLNDLSLTREDEQIEFLRLQVLEQQHVIDDLS-LEBERLLRSKRHRGKS 539
QY 451 TTPDMMEMTNGAGGEYSSENSITVAQVQYAYDNLIRINNILEDLSKAWCREQHRALV 510
Db 540 LKPKKHVWETFFGFDSESDSETLSYNTDRTDTPATPEEDLDETTTREET---ADLR 597
QY 511 WNELSKINPTSVMSMIYRNPVSIAKRGDIVSVNCIVVDQTSVLSHLKSLLSASDEKCF 570
Db 598 FCQLTR--EYQALQRAY--ALLOEQVGGTLDAR--EARTREQLQADLLRCQAKTIED-- 648
QY 571 SRPPTFKFNDSTIYKQGLGVNNEILLTTTYLETQENTEYVYFOAKTDMYIYKNVYHLK 630
Db 649 -----LEKLLVEKGQI-----YLEMEEN-----QLKSEMQADKQNEL- 682
QY 631 TVPLSSITLDTFTALNFTLLENVDFKVELYTRDEKRLSNVFDIE-TMFPREYNYAQRV 699
Db 683 -----LEFRVLEERE--RRSPAFNLQITTPENNSSALQL 717
QY 690 -----SGLRKOLLDLSTNRNQFVDAFGSLMDLGL-----AVQTVVNAV 728
Db 718 FCHQEGVWKLKQIEGTAAULTQKMDLEKBDQFSROKGYLEQLEQELDYRKALDQAYLKIQ 777
QY 729 SGVATLPSI-----VTGTFNFKNPFPGMLIIVIGVLFAYIFL 769

Db 778 DLEATLYNALQOEPGRRASEALSASQEDLQAQAVEKVRQ-----L 818
QY 770 TKTKIYETAPIK-----MIYPEIDKLKREGKSEIAPISEEBELERIV 812
Db 819 LRQSRFDSQILRRERMELLQAQOKKIREMEGKLELQORQLKELEKV 865

RESULT 37
US-10-679-063-18166
; Sequence 18166, Application US/10679063
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52054)B
; CURRENT APPLICATION NUMBER: US/10/679,063
; CURRENT FILING DATE: 2003-10-02
; PRIOR APPLICATION NUMBER: 60/415,758
; PRIOR FILING DATE: 2002-10-02
; NUMBER OF SEQ ID NOS: 27373
; SEQ ID NO 18166
; LENGTH: 889
; TYPE: PRT
; ORGANISM: Plasmodium falciparum 3D7
US-10-679-063-18166

Query Match 2.6%; Score 117; DB 6; Length 889;
Best Local Similarity 19.5%; Pred. No. 28;
Matches 151; Conservative 118; Mismatches 303; Indels 204; Gaps 35;

QY 185 TGGNLLTYVD-----RDDINMTVFLQPVGDVTPDVKRYG-----SQPELYLEPGWFGS 233
Db 99 TGGSLSLDDLYLDLRKEQKYTWMTVTKGVTGP-RRIGHVWVYSKPNLIV---FGGN 153
QY 234 YRRRTTNCNCELMDFARSNPPDFVTATGDTVEMSP-----FWS-----G 274
Db 154 DGQNTLNDVWYMHV---EMPPFEWVRVLIIPNTCKVPPQVRVHSADMCKECPASGMIVIFG 210
QY 275 EDDHENKMKHPWFVSVINNYK---VDYQNRGTVPGLG---TRIFLDREYTLSEKHL 328
Db 211 GRSANENKSLDDTWGLURQHRDGRWDWEAPIKGSGPPEARYQHTSVFIGSKIFLGG--- 266
QY 329 KMSYCPLTLWKAFYNGIQTEHSGSVHFVANDITASFTTSKEDMKEFNTTYHCLNEIEKA 388
Db 267 RNDGCAVPLSTALYN-----TTIEW 288
QY 389 EIEKKYAKVNSTHSGYGLKY-FKTDGGLYLVMOPLIQN-----RLDLAKNKLNE- 438
Db 289 VTLPSISIKFRHTSWVY---KYTIYTFGGFSSHQTOQVPTNELECELCFNLLSLNSLDSEK 345
QY 439 --TYSRRSRQAESITDPMWMTG---NGAGEYSENSITVAQVQYAYDNLIRINNIL 493
Db 346 KKSIIKOSLSLKQLTNDNLKHSDDLNRVNSYNLNSQDVINTQQHNIISTNN-QPNVSNEL 404
QY 494 EDL--SKAWCREQHRALVWNLKSNLNPVSIAKRGDIVSVNSNCIVVDOT 551
Db 405 YDLKNNASICNTLANVPIVPNVQNVNPNVNTHYRNMFDTSNNSVSRFLSRNPSNKI 464
QY 552 SVSLHKLRLLSASDEKCFSRPPTVKFMNDSTIYKQGLGVNNEILLTTTYLETQENTE 611
Db 465 RLSAHAHAHQENGSDPAFLVR---KISIDKLEEBGR-KINNGVLCTPWNYSIEFKNV 518
QY 612 YFQAKT-----DMYIYKNVYHLKTVPLSSITLDTFTALNFTLLENVDFKVELY 662
Db 519 YDKIITLLNPNTQEIQIYNHNSSEIFIPWANISVLCSI-----VIDIF 564
QY 663 TRDEKRL-----SNVFDIETMFPREYNYAQRVSGLRKOLL-DLSTNRNQFVD 708
Db 565 KQEDMWLKLRAPIKIYGDHIGQYDLMRMFOLYKCPVEEDLGEKNAIGDIDSDNDFLFG 624
QY 709 AFGSLMDDLGAQGQTVVNAVSGVATLFSSTVGTFINFKNPFPGMLI-----VWIGV 762
Db 625 DY-----VDRGS-----NSLEVICLLFA-----LKCKYPKQIHLIRGNHEDVAINS 665

```
QY 763 LFAIYFLTKK-----TKIYETAPIKMIYPEIDKL---KEREGKS--BIA 801
Db 666 LYGOBECKRLKEDVTDKDCWYQINQVFWELPICAIVE--DKILCVHGIGKGINQIS 723
QY 802 PISEELERIVLAMIHQONSHWEIKTR--KDP--KDSILTTRAONMLKRSYCNL 853
Db 724 DIS--QLKRPVVSPQNLNEQKVTDLLWSPTDNDSTLGTIPNDIRPDGTGHI 777

RESULT 38
US-10-205-516A-18
; Sequence 18, Application US/10205516A
; GENERAL INFORMATION:
; APPLICANT: Zhong, Jun
; TITLE OF INVENTION: A Method For Producing Biologically Active Botulinum Neurotoxins
; FILE REFERENCE: JZbtxl
; CURRENT APPLICATION NUMBER: US/10/205,516A
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18
; LENGTH: 1301
; TYPE: PRT
; ORGANISM: Clostridium botulinum
US-10-205-516A-18

Query Match 2.6%; Score 116; DB 6; Length 1301;
Best Local Similarity 20.8%; Pred. No. 45;
Matches 117; Conservative 78; Mismatches 194; Indels 174; Gaps 31;

QY 257 FVVTATGDTVMSPFWS--GEDDH--ENKMHKEPWFVSVINNYK--VVDYQNRGTVPGLGTR 312
Db 256 FFMQST-DTIOAEELYTFGGQDPSIISPTDSKSIYDKVLQNRGIVDRINKVLVCISDPN 314
QY 313 IFLDREYTLSEKHLKMSYCPCLTKWAFYNGIQTEHSGSVHFVANDITASFSTSKEDM 372
Db 315 INI-----NIYKNKFKDKYKFV--EDSEGGKYSI---DV 342
QY 373 KEFNTTYHCLN--BEIKAEIEKKYKVNSTHSGYDLKYPKTDGGLVNVQPLQNRL 429
Db 343 ESNFKLYKSLMLGFTET--NTAENY-KJKTASYFSDS-----LPPVK--- 382
QY 430 DAKNKLNNETYSRRRROAESTTDPMMEMTNGAGGEYSSENSITVAQVAYDNLRI 489
Db 383 -IKNLLDNEIY-----TIEGFMISDKNMCKEYRGQNKANKQ---AY----- 421
QY 490 NNILEDLSKAMCREOHRAALVWNLKINPTSVMSMIYNRPV--SAKRIGDIVSVNSCIV 547
Db 422 ---BEISK-----BHLA-----VYKIQMCKSVKIEGRVPGI--CID 452
QY 548 VDOTSVSLHLSRLLSASDEKCFSS-----RPPVTFKFMNDSTIYKGLGVNNEILLTT-- 600
Db 453 VNEN-----LFFIAKDSFSDDLKSKNERVEYNTQNN-----YIGNDFPINELIFD 501
QY 601 -TYLETQENTYFYFQAKTDMYIKNYEHLKTVPLSSITLDTFIALNFTLLENVDFKVI 659
Db 502 ISKIELPSENTESLDFNVDPVYKQPAIKKVFDENTIYOYLYSOTFPL-----NIRDI 557
QY 660 ELYTRDEKRL---SNVFDIETMFRYNYAQRV-----SGLRKDOLL--LSTNRNOFV 707
Db 558 SLTSSFDALLVSSKVSYSFFSM--DYIKTANKVVEAGLFAGWVKQIVDDFVIEANKSSTM 615
QY 708 DAFGSL-----MDDLGAQGVTVNNAVSGVATLFSISIVTGFNFINKPFGGMLM 755
Db 616 DKIAISIVPIYIGALNVGDBTAGK----NPFESAIEIAGSSILLEFI-----PEL 662
QY 756 ITWIGVLFAYIYFLTKKTIYET 778
Db 663 LIPVGVFLLESYINDKNKIYKT 685

RESULT 39
```

```
PCT-US03-27401-316
; Sequence 316, Application PC/TUS0327401
; GENERAL INFORMATION:
; APPLICANT: TUFTS UNIVERSITY
; TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE FOR DIAGNOSIS, TREATMENT AND
; FILE REFERENCE: 700355-52941-PCT
; CURRENT APPLICATION NUMBER: PCT/US03/27401
; CURRENT FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US 60/407,082
; PRIOR FILING DATE: 2002-08-30
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 316
; LENGTH: 1881
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
PCT-US03-27401-316
```

```
Query Match 2.6%; Score 116; DB 1; Length 1881;
Best Local Similarity 19.3%; Pred. No. 63;
Matches 95; Conservative 89; Mismatches 191; Indels 118; Gaps 24;

QY 432 KKNLNNETYSRRRROAESTTDP-----PMEMTNGAGGEYSSE-----NSIT 474
Db 393 KSELNNQI---DKASSVSPDTYSTASYNALGPVLET---AKGVYASEPVKQPEVNSE 444
QY 475 VAQVQYADNLRI---RINNILEDLSKAWCREOHRAALVWNLKINPTSVMSMIYNRPV 531
Db 445 -NKLTATDALNVKDTLNNTIAD-AKTKVKE-HYSDRSWQNL-QTEVTYKAEKVAANTDA 500
QY 532 SAKRIGDIVSVNSCIVVDQTSVLHLSRLLSASDEKCFSPVPV---TFKFMNDSTI--- 585
Db 501 KQSEWNAEVEKLTIATIEKLVELS-EKPILTLTSTDKILEREAVAKYTLNQNKTKIKSI 559
QY 586 ----YKQQLGVNNBIL---LTTYLETQENTYFYFQ---AKTDMYIKNYEHLKTVPL 634
Db 560 TAECLKGEEVINTVLTDDKVTTETISAAFPKNLYEYKEYTLSTTMIYDRNGEETETLEN 619
QY 635 SSITLTDTFALNFTLLENVDFKIELYTRDEKLSNVFDIETM-FREYNYAQRVSGLR 693
Db 620 QN-----IQDLUKKVELKNIKRTDLIKYENGKETNESLTTIPDDKSNYYLKITSNNQ 672
QY 694 KDLDLSTNRNQFVDAGSLMDDLGAQGVTVNNAVSGVATLFSISIVTGFNFINKPFGCM 753
Db 673 KTTI-LAVKNIETTVNGTPVYKVTADNLVS-----RTADNKFEEYVHYIEKP----- 722
QY 754 LMIIVIGVLFAYIYFLTKTIYETAPIKMIYPEIDKLKE-----REGKS-----E 799
Db 723 -----KVHE---DNVYVNFKELVEAIONDPSKEYRLGQMSARN 758
QY 800 IAPISEELERIVLAMIHQONSHMETKTRKPKDSILTTRA-----QNMLKRSYCN 852
Db 759 VVPNGKSYITKEFTGKLJLSEGGKQFAITELEHPLFNVTNATINNVAENFENVEIERSQDN 818
QY 853 LKNAESVEMNLTL 865
Db 819 I-----ASLANTM 826
```

```
RESULT 40
US-10-472-928-1236
; Sequence 1236, Application US/10472928
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026926WO
; CURRENT APPLICATION NUMBER: US/10/472,928
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: GB-0107658.7
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 4979
```

RESULT 42
US-10-472-928-4712
; Sequence 4712, Application US/10472928
; GENERAL INFORMATION:
; APPLICANT: CHIRON SEA
; APPLICANT: THE INSTITUTE FOR GENOMIC RESEARCH
; TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026926WO
; CURRENT APPLICATION NUMBER: US/10/472,928
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: GB-0107658.7

Db 1354 --DLV--RCI--RNSLVHKKDI-LRSITQQQ--NLPPSKFEPANDCL-----1391
Qy 596 ILLTTTILETCOENTYEFQAKTDMYIKYKYEHLKTVPLSS 636
Db 1392 -----KQESVLLYFECKFPDLLFHLHYQQPKSDFNS 1423
RESULT 44
PCT-US03-06962-40
; Sequence 40, Application PC/TUS0306962
; GENERAL INFORMATION:
; APPLICANT: The Forsyth Institute
; TITLE OF INVENTION: Immunogenicity of Glucan Binding Protein
; FILE REFERENCE: 25669-018-061
; CURRENT APPLICATION NUMBER: PCT/US03/06962
; CURRENT FILING DATE: 2003-07-01
; PRIOR APPLICATION NUMBER: 10/383,930
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: 60/402,483
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 60/363,209
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 40
; LENGTH: 1518
; TYPE: PRT
; ORGANISM: Streptococcus salivarius
PCT-US03-06962-40
Query Match 2.5%; Score 113.5; DB 1; Length 1518;
Best Local Similarity 18.1%; Pred. No. 71;
Matches 146; Conservative 116; Mismatches 296; Indels 247; Gaps 35;
Qy 135 YRKVVTTSTVNGYISDITNQHTFYKSIEP-----WETEKMDTI 174
Db 393 YRLNRTATNQTGTIDKSILDEQS-----DNHMGGFDFLLANDVDLSNPVQAEQLNQI 447
Qy 175 YQCNSLRNLGNGLLTYVDRDDINMTVFLOPQDGVTPDKVYGSQPELYLEFGWFGSY 234
Db 448 HYLNMWGSIVMG-----DKDANFDGIRDAVDNVDADM-----LQLYTNYF-----488
Qy 235 RRRRTTVCCELMDMFARSNPDPFVTATGDTVEMSPFWSGEDHEN-----KHEKP 286
Db 489 REYGVNKKSEANALAH-----ISVLEAWSLNDHNDYNDKTDGAALAMENKQ 533
Qy 287 WFSVINNYKVVDYQNRGTVPGLKTRIFLDREYTLTW-----EKHLKMSYCPPLTWK 340
Db 534 RLALLFSLAKPIKERTPAVSPLYNNTTQDRDEKTDWINKDGSKAYNEDGTVKQSTIGK 593
Qy 341 AFYNGIOTEHSGSYHFV-----ANDITA-----SFTSKEDMKENFTTYHCL 382
Db 594 --YNEKYGDASGNVYFIRAHNNVQDIIAETIKKEINPKSDGFTTIDAEMKQ---APELY 648
Qy 383 NEETKAEIEKKYAKVN-----STHSKYGDL-----KYFKTDGGLYLVWQPL 423
Db 649 NKDWLSS--DKRYTLNNIPAAVAVMLQNMETIRVYVYGLDYTDGHHYMETKSPYDITLVN 707
Qy 424 IQNRL-----LDANKLNN---ETYRRRRRQAEESTTDPMMEMTNGAGGEY 467
Db 708 MKSRIVKVGSGQAQSRVYLPDTRGWDNSDVELY--RTNEVYTSVRYKDIWANDTEG--763
Qy 468 SSENSITVAQVQAYDNLRIINNILE---DLSKAWCREQRAALVWNLKSKINPTSVMS 524
Db 764 -SKYSRTSGQVTLVANNPKMLDOSAKLVNEMGXIHANQYRALIVG-----809
Qy 525 MIYNRPVSARKRIGDIVISVNCIVVDQTSVLHSLKRLLSASDEKCFSPRPVTFKFMNDST 584
Db 810 -----TAGIKNFTSDADAIAAGYKVTEDNSGVLTFGANDIKGYE-----TFDMSGFVA 858
Qy 585 IYKQGLGVNNEILLTTTYLETCOEN-----TEYFQAKTDMYIKYKYEHLKTVPLSSIT 638
Db 859 VWVPVGASDQDIRVAPSTEAKKBEGLTLKATEAY---DSQLIYEGFSNFQITPDGS--912

Qy 639 TLDTFIALNFTLLENVD-----FKVIELYTR--DEKRLSNVFDIETMPREYNVYAO 687
Db 913 --DPSYTTNRKTAENVDLFPKSGCVTSFEMAPQFVSADGDTFLDSVIQ-----NGYA- 961
Qy 688 RVSGLRKDLIDLSTNRNQFVDFAGFSLMDDLGAVGQTVVNAVS-----GVATLFS 736
Db 962 -----FADRYDLAMSKNNKYGSKEDLRDALKALHAGIAIADWVPDQIYQLPGKEVVA 1016
Qy 737 SVITGFINKPFGGMLMIIVIGLVFAIYFUTKKT--KIYET-----APIKMIYPE 787
Db 1017 TRTDG-----AGRKIADAIID--HSLYVANSKSGKDYQAKYGGEFALBKKAYPE 1065
Qy 788 I-----DKLKEREGKSE 799
Db 1066 MFKVNMISTGKPIDDSVKLKQWKA 1090
RESULT 45
US-10-679-063-18152
; Sequence 18152, Application US/10679063
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52054)B
; CURRENT APPLICATION NUMBER: US/10/679,063
; CURRENT FILING DATE: 2003-10-02
; PRIOR APPLICATION NUMBER: 60/415,758
; PRIOR FILING DATE: 2002-10-02
; NUMBER OF SEQ ID NOS: 27373
; SEQ ID NO 18152
; LENGTH: 1617
; TYPE: PRT
; ORGANISM: Plasmodium falciparum 3D7
US-10-679-063-18152
Query Match 2.5%; Score 113.5; DB 6; Length 1617;
Best Local Similarity 18.0%; Pred. No. 75;
Matches 105; Conservative 84; Mismatches 193; Indels 201; Gaps 25;
Qy 267 EMSPFWSGE---DDEHNKM-----HEKPNF-----VSVINNYKVVDYQNRGTV 306
Db 511 ELFSFGNGKLYNREENKIEVATFGCKNKNWFEKYDVKIENCKYNNYKNENNNDKNKN 570
Qy 307 PL---GKTRIFLDREYTLISWEKHLKMSYCPPLTWKAFYNGIOTEHSGSYHFVANDITA 363
Db 571 PGFKGLDHIKEEYD---KMYMDGNNYMKGLIGNISPKTORSHD-----NNI-E 620
Qy 364 SFTTSKEDMKERN-----TTHCLNEEIKAE-----IEKKYAKVNSTHSGYDLY 409
Db 621 NYISDSSENWNTYIMNEVNRVLPYNNINNNNNNNNNKLIYNQGTLSIEHTNNISKY 680
Qy 410 FKTGGLYLVWQPLIQNRLLDAKNLNNETYSRRRQAEESTTDPMMEMTNGAGGEYSS 469
Db 681 TTVD-----YSNDIL---KTGKTYKELTKRKEEICELL-----NSEDKK 718
Qy 470 ENSI-----TVAQVQAYDNLRIINNIL-----ED 495
Db 719 KHNKHKFLDKMKSKKHIEDFFKEDKICERSEKKEKLIKFTDILHAGAFGHFHRNKG 778
Qy 496 LSKAWCREQRAALVWNLKSKINPTSVMSMIYNRPVSARKRIGDIVISVNCIVVDQTSVL 555
Db 779 LRSFVSSDVKRNNSNSINKKKDQLK--YN-----TL 811
Qy 556 HKSRLRLLSASDEKCFSPRPVTFKFMNDSTIYKQGLGVNNEILLTTTYLETCOENTYEFQ 615
Db 812 EKYETKIGAFDESKLG-----NVRNDKTVSYNQIECKNSI-----NYN 849
Qy 616 AKTDMYIYK-----NYEHLKATVPPLSSITTLDTFIALNFTLLENVDVFKVIELYTR--DEKRL 669
Db 850 SKTELNIDKNKDFNFKNKWKYKMNDEEKIHLKFLPNGLDNDKKELKSNNDGIDDKLL 909
Qy 670 SNVFDIETMFREYNVYQVSVGLRKOLL-----DLSTNRNQFVDFAGSL 713

Db	910	SYVEDVK--YKYNDSKQKNGIPKQIVNNRRNMINKDYEEKYNIKVNRN-----	958
QY	714	MDDLGAVGQTVVNAVSGVATLFSSIVTGFINFKNPF--GGML 754	
Db	959	-----NEENGVDKMLSVHGGVFN--KNDFEYGGKI 987	

Search completed: October 28, 2003, 15:40:25
Job time : 44 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 15:27:11 ; Search time 47 Seconds
(without alignments)
1769.913 Million cell updates/sec

Title: US-10-055-364-24

Perfect score: 4547

Sequence: 1 MAGSLKRGSLWLYQV.....KRSGLYNLKNVSEMLNTL 865

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_76: *
1: PIR1: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2489.5	54.8	854	2 T03107	glycoprotein B - a
2	1953.5	43.0	874	2 S5602	glycoprotein B - e
3	1949.5	42.9	808	1 VGBESM	glycoprotein B - pre
4	1925	42.3	807	2 T42924	glycoprotein B - a
5	1848.5	40.7	857	1 Q0BE1L	glycoprotein B - h
6	1787	39.3	874	2 S25530	glycoprotein B - b
7	1128	24.8	907	1 VGBETE	glycoprotein B - pre
8	1121.5	24.7	906	1 VGBEC1	glycoprotein B - pre
9	1084.5	23.9	822	2 T41941	glycoprotein B - h
10	1052.5	23.1	830	2 T43999	glycoprotein B - lim
11	1050.5	23.1	830	1 A44047	glycoprotein B - pre
12	1041.5	22.9	830	1 B44047	glycoprotein B - pre
13	1040.5	22.9	830	2 T44186	probable glycoprotein B - pre
14	1039	22.9	831	1 VGBE6S	glycoprotein B - h
15	1034	22.7	928	1 VGBEMC	glycoprotein B - pre
16	894	19.7	920	1 D48349	glycoprotein B - pre
17	876.5	19.3	865	1 VGBEBR	glycoprotein B - pre
18	862	19.0	948	2 A56602	glycoprotein B - hom
19	857	18.8	932	1 VGBEBB	glycoprotein B - pre
20	845.5	18.6	903	1 VGBEB1	glycoprotein B - pre
21	845.5	18.6	903	1 VGBEK1	glycoprotein B - pre
22	841	18.5	904	1 VGBEW7	glycoprotein B - pre
23	839	18.5	904	1 VGBEB2	glycoprotein B - pre
24	837	18.4	904	1 VGBEK2	glycoprotein B - pre
25	836	18.4	883	1 VGBEIS	glycoprotein B - pre
26	835	18.4	980	1 VGBEC6	glycoprotein B - pre
27	833	18.3	980	1 VGBEC5	glycoprotein B - pre
28	832	18.3	868	1 VGBE31	glycoprotein B - h
29	830	18.3	883	1 VGBEIL	glycoprotein B - pre

glycoprotein B - i
glycoprotein B pre
glycoprotein B pre
glycoprotein B pre
probable envelope
glycoprotein B pre
glycoprotein B - f
glycoprotein 14 pr
glycoprotein g1i p
glycoprotein gi pr
surface layer prot
ORF MSV077 hypothe
rhopty protein -
hypothetical prote
p101 protein precu
reticulocyte-bind

ALIGNMENTS

RESULT 1

T03107

glycoprotein B - alcelaphine herpesvirus 1

C:Species: alcelaphine herpesvirus 1

C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 24-Nov-1999

C:Accession: T03107

R:Ensser, A.; Pflanz, R.; Fleckenstein, B.

J. Virol. 71, 6517-6525, 1997

A:Title: Primary structure of the alcelaphine herpesvirus 1 genome.

A:Reference number: 214840; MUID:97404659; PMID:9261371

A:Accession: T03107

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-854 <ENS>

A:Cross-references: EMBL:AF005370; NID:G2337967; PIDN:AACS8059.1; PID:G2337975

C:Superfamily: herpesvirus glycoprotein B

C:Keywords: glycoprotein

Query Match 54.8%; Score 2489.5; DB 2; Length 854;

Best Local Similarity 57.0%; Pred. No. 5.3e-140;

Matches 471; Conservative 135; Mismatches 173; Indels 47; Gaps 8;

Qy	64	NENRDSSEQN---	KNYIGSPSTPEYRVCSAGVGDFVFOFDHVCPPASDMVHSEGIILLI	120
Db	42	SDNIIRQNNNTAKGIHSDPSAFPRVCSASNIGDIFRFQSHSCPNTKQKEHNEGILLI	101	
Qy	121	YKQNIIPFMFRVRYKRVKVTITSTVYNGIYSDSIITNQHTFYKSIIEPWETEKMDTIYQCFNS	180	
Db	102	FKENIVPVFKVRYKRVKVTITSTIYNGIYADAVTNQHVFSKSVPIYETRRMDTIYQCNS	161	
Qy	181	LRLTNGNLLTYVDRDDINMTVFLQPDGVTPDVKRVGSOPELYLEPGWFGVSRRTTV	240	
Db	162	LDVTVGNNLLVYTDNGSNMTVDLPQDGLSNSVRRYHVSQPEIHAEPGWLLGGYRRRTTV	221	
Qy	241	NCELMDMFARSNPDEFVATGDTVMSPPW-----SGEDDHENKMKHEKPFVSV	291	
Db	222	NCEYETDARAVPPFRFYITNIGTIEMSPWMSKAWNETEFSGEPRDT-----LTV	272	
Qy	292	INNYKVVYDQNRGVPLGKTRIFLDRBEYTLISWEKHLKNMSYCYPLTLWAFYNGIQTEHS	351	
Db	273	AKDYRVVDYKRGTPQPOGHTRFIVDKEEYTLISWQAQFNISYCRWAHWKSFDAIKTEHG	332	
Qy	352	GSYHFVANDITASPTTSKEDMKENFTTYHCLNEIEIKAEIEKKYAKVNSTHYSKYGDLYKFK	411	
Db	333	KSLHFVANDITASPYTPNTQTRVLGKHVCLNNTIESELKSRSLAKVNDTHSPNGTAQVYL	392	
Qy	412	TDGGLYLVQPLIQNRLLDKAKNKL-----NETYSRSRRQAESTTDPMMEMTNG	462	
Db	393	TNGGLLVQPLVQVQKLLDAKGLLDVAKKQONTTTTTTTTTSRRQRQSVS-----SG	444	
Qy	463	AGGYSSSENSITVAQVOYAYDNLIRINNILEDLSKAWCREOHRALVWNELSKINPTSV	522	

445 IDVVYTAESTILLTQIOFAYDITLRAQINNVLLEELSRACREOHRASLMNLSKINPTSV 504
523 MSMLYNRPVSAKRIGDVISVNCIIVDQTSVLSLHKSRLLSA-SDEKCFRPPVTPKPMN 581
505 MSSYIGRPVSAKRIGDVISVSHCVVDQDSVLSRSMRVPGDKTHCEYCRPPVTPKFIN 564
582 DSTYKGGQGVNNELLTTTLETQCENTEYFOAKTDMYIKNYEHLKTVPLSSITLTD 641
565 DSHLYKQGVNNELLTTTAVELCHENTHYFOGGNNMYFYKNYRAVKTPMPGDVATLD 624
642 TFIALNFTLLJENVDFKVIYTRDEKLSNVFDIETMFREYNYAQRVSGLRKDLDLST 701
625 TFMVNLTLVENIDFVIELYSREKRMSTAFDIETMFREYNYTQRTVGLRDLTDLAT 684
702 NRNOFVDAFSLMDDLGAQGVTVNAVSGVATLTFSSIVTGFIFIKNPFGLMLIIVIG 761
685 NRNOFVDAFSLMDDLGVGVTVNAVSSVATLTFSSIVSGIINFIKPNPFGMLLFGIIAA 744
762 VLFAYIFLTKTKIYETAPIKMIYPEIDKLEREGKSEIAPISEEELERIVLAMHIHOON 821
745 VVITVILLNKAKEFAQNPQVMYIPDKITTSQREELQVDPISKHELDRLMLAHDIHAS 804
822 SHMETK-----TRDKPKOSILTRAQNMRLKRSYGNLKNABSE 860
805 KPESKQDEOGSTTSGPAD-WLNKAKNVLRRRAGYKPLKRTDSFE 849

RESULT 2
S55602
glycoprotein B - equine herpesvirus 2
C:Species: equine herpesvirus 2
C:Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 26-Aug-1999
C:Accession: S55602
R:Telford, E.A.R.; Watson, M.S.; Aird, H.C.; Perry, J.; Davison, A.J.
J. Mol. Biol. 249, 520-528, 1995
A:Title: The DNA sequence of equine herpesvirus 2.
A:Reference number: S55594; MUID:95302501; PMID:7783207
A:Accession: S55602
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-874 <TEL>
A:Cross-references: GB:U20824; NID:G695172; PIDN:AA13795.1; PID:G695180
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1995
A:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein

Query Match 43.0%; Score 1953.5; DB 2; Length 874;
Best Local Similarity 46.1%; Pred. No. 3.8e-108;
Matches 403; Conservative 147; Mismatches 250; Indels 75; Gaps 18;

11 VLALWLYQVALYSLSTAETGVTPPNATWSTPESPLTGHYGHDSHSGRGNNENRDSF 70
10 VLCLWCY--AALLCQGAQEWVA-----ETTP----FATH-----RPEVAAE 46
71 EONKNYIGSPSTFPYRVCASGV-GDVRFQTDHVCPCDASDMVHSEGLLIYKONIIPFM 129
47 ENPANPF-----LDFVCGASPTGGEIIFRPLESCNTEKDKHIEGLIYKNTIVPVY 101
130 FVRKRYKRVVTTSTVYNGIYSDSITNHTYKSIETPETERKMDTIYQCFNSLRINTGNNL 189
102 FNRKRYKRMKMTSTTIYKGSSEDAITNQTSTRSYAVPLVEVQMDHYQCFSAVQVNEGHV 161
190 LTYVDRDDINMTVFLQPDVGPDPVKYQSGOPELYLEPGFWGWSYRRRTTVNCELMDMFA 249
162 NTYYDRDGMNETAFKPKADGLTSSITRYQSQOPEYATPRNLLSYTRITTVNCEVTEMSA 221
250 RSNPFPFFVATGDTVMGSPFWSGDEDDHENKMKHKEPWSVNNYKVDYQNRGTVPGLG 309
222 RSMKPFPPVTSVGDITTEMSPFLKENGTEPEKILKRPHSLOLKNVATYK-----GVGLG 277
310 K-----TRIFLDREBYTSLWEKHLKNMSYCPPLTLWKAIFYNGIQTSEHSGSYHFVANDITASF 365
478 QADNATFFAIFGYSLSWKATTENSSYCDLILWKGFNSAIQTQHNSSLFIANDITASF 337

366 TTSKEDMKBENTTYHCLNBEIKAEIEKKYAKYNSTHYSKYGDLKYFKYKTDGGLYLWMOPLIO 425
338 STPLEBEANFNETFKCIWNNTQEEIOKKLUKEVKEKTRPNTGAKVYKTKTGLNYLYWMOPLIO 397
426 NRLDDAKNLNN-----ETYSRRSRQAEASTTDPMMEMTNGAGGEYSS-ENS 472
398 IDLLDTHAKLYNLNATASPTSTPTTSPPRRRDTSS-----VSGGNGNGDSTKEES 450
473 ITVAQVQAYDNLRIIRINNILEDLSKAWCREQHRALVNWELSKINPTSVMSMIYNRPVS 532
451 VAASQVQFAYDNLRLKSRINRVGLSELRACREQYRASLMWYELSKINPTSVMSAIYGRPVS 510
533 AKRIGDIVSVNCIIVVDQTSVLSLHKSRLLSASDEKCFRPPVTPVTFKMDNSTIYKKGOLGV 592
511 AKLIGDVSVSDCISVDQKSVFVHKNMK-VPGKEDLCYTRPVVGVKPFINGSELSFAGOLGP 569
593 NNEILLTYYLETQCENTEYFOAKTDMYIKNYEHLKTVPLSSITLDTFFIALNFTLLE 652
570 RNEIVLSTQVEVCOHSCHEHYFQAGNQMYKDYVYSTNLNTDITPLHTMITNLNLSLVE 629
653 NVDFKVIYLYTRDEKLSNVFDIETMFREYNYAQRVSGLRKDLDD-LSTNRNQFVDAPG 711
630 NIDFKVIELYSKTEKLSNVFDIETMFREYNYTQNLNGLRKDLDDSIDHGRDSFIQTLLG 689
712 SLMDLGAQGVTVNAVSGVATLTFSSIVTGFIFIKNPFGLMLIIVVI-GVLFAIYFLT 770
690 DIMQDLGTGKVVVNVNAVSGVSLFGSIVSGVIFSPFNPFGLMLIIVLIITAGVVVVYLFMT 749
771 KTKIYETAPIKMIYPEIDKLEREGKSEIAPISEEELERIVLAMHIHOONSHMETKTRK 830
750 RRSRIY-SAPIRMLYPGVERAAQEPCH--PVSEDOIRNLMGMHGFQORQRAEBAER 805
831 DPK-----DSILTRAQNMRLKR---SGYSLNK 854
806 EBEVKGKRTLFEVIRDSATSVLRRRRGGGGYQRLQ 840

RESULT 3
VBGESM
glycoprotein B precursor - saimirine herpesvirus 1 (strain 11)
C:Species: saimirine herpesvirus 1
C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 16-Jul-1999
C:Accession: D34126
R:Albrecht, J.C.; Fleckenstein, B.
Virology 174, 533-542, 1990
A:Title: Structural organization of the conserved gene block of Herpesvirus saimiri codir
A:Reference number: A34126; MUID:90163221; PMID:2154888
A:Accession: D34126
A:Molecule type: DNA
A:Residues: 1-808 <ALB>
A:Cross-references: GB:M31122; NID:G331052; PIDN:AAA46164.1; PID:G331056
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein; transmembrane protein
F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-808/Product: glycoprotein B #status predicted <GLB>
F:675-692/Domain: transmembrane #status predicted <TM1>
F:702-722/Domain: transmembrane #status predicted <TM2>
F:30,158,239,251,285,331,344,355,361,471,532,569,587,598,727/Binding site: carbohydrate

Query Match 42.9%; Score 1949.5; DB 1; Length 808;
Best Local Similarity 46.6%; Pred. No. 5.8e-108;
Matches 373; Conservative 159; Mismatches 236; Indels 33; Gaps 10;

71 EONKNYIGSPSTFPYRVCASGVGVDFRFTQDTHVCPCDASDMVHSEGLLIYKONIIPMF 130
28 EKNKTOAIYQEVFKYRVCASGASITGELFRFDLORTCPSTEDKVHKEGILLVYKKNIVPYIF 87
131 RVRKRYKRVVTTSTVYNGIYSD--SITNQHTFYKSIETPETERKMDTIYQCFNSLRINTGNN 188
88 KVRKYKRTTIVRIEIRREGVATNKNWELSRAPVKYEIDIMDKTKYQCHNCMQIEVNGM 147
189 LLTYVDRDDINMTVFLQPDVGPDPVKYQSGOPELYLEPGFWGWSYRRRTTVNCELMDMF 248

A;Contents: annotation; protein coding region
C;Keyfamily: herpesvirus glycoprotein B
C;Keywords: glycoprotein; transmembrane protein
F:76.163.290.329.348.395.436.563.629/Binding site: carbohydrate (Asn)
#status (covalent)

[illegible]

RESULT 6
S25530 glycoprotein B - bovine herpesvirus 4
C:Species: bovine herpesvirus 4
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999
C:Accession: S25530
R:Goltz, M.; Broll, H.; Gielow, A.; Weigelt, W.; Ludwig, H.; Borchers, K.

submitted to the EMBL Data Library, August 1992

A;Description: Glycoprotein B of bovine herpesvirus type 4.
A;Reference number: S25530
A;Accession: S25530
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-874 <GOL>
A;Cross-references: EMBL:Z15044; NID:G58689; PID:G58690
C;Superfamily: herpesvirus glycoprotein B
C;Keywords: glycoprotein

	Query Match	39.3%	Score 1787;	DB 2;	Length 874;
	Best Local Similarity	43.5%;	Pred. No. 2.9e-98;		
	Matches	367;	Conservative 159;	Mismatches 248;	Indels 70; Gaps 20;
Qy	34	SPNTATWSTESPLTGHYCTHDSHGERGNNENRDSEEQNKNIYGPSSTFPYRVCSA-SG	92		
Db	58	SPONT---STSKPSTONQGTSTPT-----IPVTDDTASKNY-----KYRVCSASS	102		
Qy	93	VGVVFRQTDHVCPCDASDMVHSEGIILLIYKONIIPFMFRVKRYKRVKVTSTVYNGIYSDS	152		
Db	103	SGELFRFDLDQCPDKOKKHVEGILLVLKKNIVPIFKVKRYKIATSVTVYRGSQAA	162		
Qy	153	ITNQHTFYKSIPEWTEKMDTIIYQCNFSRLNTGGNLLIYVDRDDINMTVFLQPVGVTP	212		
Db	163	VTVNRDDISRAIYNEISMTIDRTVHCFESAMATVINGILNTIYIDRSENKSVLPQVAGLTE	222		
Qy	213	DVKRYSQPELYLEPCWFGWSYRRRTIVNCELMDMFARNPPDPFVATGDTGVEMSPFW	272		
Db	223	NINRYFSQPLIYAEPGWFGPIYRVRTVNCVEVDMYARSVEPTHETALGDTIEISPPFC	282		
Qy	273	SGEDDHENKM-----HEKPFVSVVINNVYDYNQNGRVPLGKTRIFLDRBEYTL	322		
Db	283	-----HNNSQCTTGNSTSRDATKVM---IEENHQTVDYERRGH-PFKDKRIFLDBEYTI	333		
Qy	323	SWEKHLKNMSYCPILTAKWAFYNGIOTHEGSGVHFVANDITASTPTSKEDMKPF--NT-TV	379		
Db	334	SWAEDREARAICDFVILWKTFPRAIQTIHNESSHFFVANEVTSAPLTSNQSETELRGNTIEL	393		
Qy	380	HCLNEIKABIEKKYAKVNSTHSGYGLKFKYKTDGGLYLVWQPLIQRLLDARK---KLN	436		
Db	394	NCMNSTINETLEETVKKFNKSHIRDCGEVYKYKINGGLFLIWOAKKPLNLSEHTNYTIERN	453		
Qy	437	NETYSRRSRQABSTTDPMMWMTGNGAGGEYSSENSITVAQVOYAYDNLRIRINNILEDL	496		
Db	454	NKTGNKSRQKRSYDT-----KTFQAGK-----LSTAQVOYAYDHLRTSMNHILEEL	500		
Qy	497	SKAWCREQRAALVNNELSKINPTSVMKSYNRPVSAGRIGDVISVNCIVVDQTSVSLH	556		
Db	501	TKTWCREQKDNUNWVLESKINPVSVMAYIYKPAVAKAMGDAFMWSECLNVDAQSVNIH	560		
Qy	557	KSURLLSASDEK-CFSRPPVTFKMNDDSTIYKQGLGVNNBEILLTTTYLTQCENTEYYPQ	615		
Db	561	KSMR---TDDPKVCYSRPLVTFKFNSTATFRGOLGTRNEILLTNTHVETCRPTADHYFF	617		
Qy	616	AKTDMWIIYKNYEHKLVTPULSSITTLDTFTALNFTLENVDVFKVIELYTRDEKELSNVFDI	675		
Db	618	VKNMTHYFKDYKPVKMTDNNISTLDTFTLNLTFIDNIDFKTVELYSYTERKMGASALDL	677		
Qy	676	ETMFREYNYAQRVSLGRKDL--LDLSTNRNQFVDAFGLSMDDLLGAGGVTVNVAVSGVA	732		
Db	678	ETMFREYNYTQKLSLRDLNLTIDL--NRDLVKDLSEMADLDGICKVVVNTSGIV	735		
Qy	733	TLFSSIVTGFINPIKNPFGGMLMIIVIGLVFAIYPLTKKTKIYETAPIKMIYPEIDKUK	792		
Db	736	TVFSGIVGGFVSFTNPIGCVTIILLIIVVVFVTVSVRRTNNMNEAPIKMIYPNIDKAS	795		
Qy	793	EREGKSEIAPISEELERIVLAMIHQOONSHMETKTRDKPKO--SILTRAQNMKLKRS	849		
Db	796	EQE---NIQPLGEEIKRILLGMHQLQOEHGKSEEAHKPGLFQILGDLQQLLRRR-G	851		
Qy	850	YSNL	853		

A:Reference number: Z22732; MUID:99412319; PMID:10482554
A:Accession: T43999
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-830 <ISE>
A:Cross-references: EMBL:AB01506; NID:94995977; PID:BAA78260.1; PID:g4996027
A:Experimental source: strain HST; pop. variant B
C:Genetics:
A:Note: U39
C:Superfamily: herpesvirus glycoprotein B

```
Query Match      23.1%; Score 1052.5; DB 2; Length 830;
Best Local Similarity 32.4%; Pred. No. 1.1e-54;
Matches 245; Conservative 157; Mismatches 295; Indels 59; Gaps 15;

Qy      83 PPRVCSAGVDVFRFQTDHVC-PDASDMVHSEGILLIYKQNIIPPMFRVRYKRVKVT 141
Db      36 YPRICSTAGTDLRMFRDRDISCSYKNAKMSGFFIYKNTIETVTPVRYKELTF 95

Qy      142 STVYNGIYSDSITNQHTFYKSIETPEWTEKMDTIYQCFNSLRN--TGGNLLTYVDRDIN 199
Db      96 QTSYRDVGYYFLDRTVMGLAMPYEAANLVNSRAQCYSAVAIKRPGDGVFSAYHEDNNKN 155

Qy      200 MTFVLPQVDGVTDPVKYGSQPELYLEPCGFWGSGYRRRTTNCCLMDMFARSNPPDFV 259
Db      156 ETLEFLPLNFKSVTKRFTITKEPYFARGPLW-LYSTSTSLNCIVTEATAKAKYPSYFA 214

Qy      260 TATGDTVMSPWMSGED-DHENKMEKHPWFVSVINNYKV--DYON--RGTVPLGKTRIFL 315
Db      215 LTTGIVGSPFPDGSNGKHFAEPLEK--LTILENTVIMDLMMNGMGATTLVRKIAFL 271

Qy      316 DREYTLTSLWEKHLKMSYCPCLTWKAFYNGIOTHSYSGVHFVANDITASFTSKEDMKEF 375
Db      272 EKGDTLFSWEIKEENESVCMKHWTTVTHGLRAETDEYHFISKELTAFAFVTPKDSINLT 331

Qy      376 NTTYHCLNEEIKAEIEKKY-AKVNSTHSGYGLKYFTDGGLYLVWQPLIQNRLDANK 434
Db      332 DPKQTCIKNEFENIIEKVMSYNDYNTYSMNGSYQIFKTTGDLILWQLVQKSLM--V 387

Qy      435 LNNETYSRRSRQAESTTDPMMEMTGAGGEYSSENSITVAQVQYAYDNLIRINNILE 494
Db      388 LEQGSVNLRRRLDLV-----DVKSRHDILYVQLQYLDYTLKDYINDALG 431

Qy      495 DLSKAWCREQRAALVWNLKINPTSVMSIYNRPVSAGKRGDVISVNCIVVDQTSVS 554
Db      432 NLAESWCLDQKRTITMLHELKISPSISVSEYGRPISAQLHGDVLAISKIEVNSQSVQ 491

Qy      555 LHKSLRLLSA-----SDEKCFSPRPVTFKPMNDS--TIYKGQGVNNEILLTTTYLETQEN 609
Db      492 LHKSMRVVDKAGLRSETVCYNRPVTFSPVNSTPEVPGQLGDLNEILLGDHRTCECIP 551

Qy      610 TEYVFOAKTDMYIKNYEHLKTVPLSSITTLDTPIALNFTLLENVDFKVIELYTRDEKRL 669
Db      552 STKIFLGNHAAVYDYTHNTSTPIEDIEVLDAFIRLKIDPLENADFKLLDLYSPDELSR 611

Qy      670 SNVFDIETMFRYNYAQRVSGRLKDLDDLSTNRNQFVDAFGSLMDDLCAVGQTVNNAVS 729
Db      612 ANVFDLENILREYNSYK---SALYTIKIAKNTPTSYNGINSFLOGGAGTGTGLGVSIS 668

Qy      730 GVATLFSSIVTGFINFKNPFGGLMIIIVIGVLFALYFLTKTKTIYETAPIKMIYP--- 786
Db      669 VTAGALGDIVGGVSWFLKNPFGGLMLILAIIVVWIIIVFVRQKHVLSKPIDMMFPYAT 728

Qy      787 -----EIDKLKREKSEIAPISEEE 807
Db      729 NPVTTVSSVTGTTAVKTPSVKADGGTSVA-VSEKE 763
```

RESULT 11

A44047
Glycoprotein B precursor - human herpesvirus 6 (strain GS)
C:Species: human herpesvirus 6
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 16-Jul-1999

C:Accession: A44047
R:Chou, S.; Marousek, G.I.
Virolgy 191, 523-528, 1992
A:Title: Homology of the envelope glycoprotein B of human herpesvirus-6 and cytomegalovirus
A:Reference number: A44047; MUID:93033157; PMID:1329336
A:Accession: A44047
A:Molecule type: DNA
A:Residues: 1-830 <CHO>
A:Cross-references: GB:M97928; NID:g325474; PID:AAA43847.1; PID:g325475
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein; transmembrane protein
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-395/Domain: Product: Glycoprotein B #status predicted <GPB>
F:24-395/Domain: gp116 #status predicted <G116>
F:400-830/Domain: gp55 #status predicted <G55>
F:656-674/Domain: transmembrane #status predicted <TM1>
F:693-710/Domain: transmembrane #status predicted <TM2>
F:155,228,247,286,329,355,361,486/Binding site: carbohydate (Asn) (covalent) #status pr

```
Query Match      23.1%; Score 1050.5; DB 1; Length 830;
Best Local Similarity 32.3%; Pred. No. 1.5e-54;
Matches 244; Conservative 160; Mismatches 293; Indels 59; Gaps 15;

Qy      83 PPRVCSAGVDVFRFQTDHVC-PDASDMVHSEGILLIYKQNIIPPMFRVRYKRVKVT 141
Db      36 YPRICSTAGTDLRMFRDRDISCSYKNAKMSGFFIYKNTIETVTPVRYKELTF 95

Qy      142 STVYNGIYSDSITNQHTFYKSIETPEWTEKMDTIYQCFNSLRN--TGGNLLTYVDRDIN 199
Db      96 QSSYRDVGYYFLDRTVMGLAMPYEAANLVNSHAQCYSAVAMKRPDGTVFSAFHEDNNKN 155

Qy      200 MTFVLPQVDGVTDPVKYGSQPELYLEPCGFWGSGYRRRTTNCCLMDMFARSNPPDFV 259
Db      156 NTLNLFPLNPKSITNKRFITTKPEYFARGPLW-LYSTSTSLNCIVTEATAKAKYPSYFA 214

Qy      260 TATGDTVMSPWMSGED-DHENKMEKHPWFVSVINNYKV--DYON--RGTVPLGKTRIFL 315
Db      215 LTTGEIVGSPFPNGSKGHFAEPLEK--LTILENTVIMDLMMNGMGATTLVRKIAFL 271

Qy      316 DREYTLTSLWEKHLKMSYCPCLTWKAFYNGIOTHSYSGVHFVANDITASFTSKEDMKEF 375
Db      272 EKADTLFSWEIKEENESVCMKHWTTVTHGLRAETDEYHFISKELTAFAVAPKESLNT 331

Qy      376 NTTYHCLNEEIKAEIEKKY-AKVNSTHSGYGLKYFTDGGLYLVWQPLIQNRLDANK 434
Db      332 DPKQTCIKDEFEKLINEVMSYNDYNTYSMNGSYQIFKTTGDLILWQLVQKSLMFL--- 388

Qy      435 LNNETYSRRSRQAESTTDPMMEMTGAGGEYSSENSITVAQVQYAYDNLIRINNILE 494
Db      389 ---EQSEKIRRRRDVV-----DVKSRHDILYVQLQYLDYTLKDYINDALG 431

Qy      495 DLSKAWCREQRAALVWNLKINPTSVMSIYNRPVSAGKRGDVISVNCIVVDQTSVS 554
Db      432 NLAESWCLDQKRTITMLHELKISPSISVSEYGRPISAQLHGDVLAISKIEVNSQSVQ 491

Qy      555 LHKSLRLLSA-----SDEKCFSPRPVTFKPMNDS--TIYKGQGVNNEILLTTTYLETQEN 609
Db      492 LHKSMRVVDKAGRVSETVCYNRPVTFSPVNSTPEVPGQLGDLNEILLGDHRTCECIP 551

Qy      610 TEYVFOAKTDMYIKNYEHLKTVPLSSITTLDTPIALNFTLLENVDFKVIELYTRDEKRL 669
Db      552 STKIFLGNHAAVYDYTHNTSTPIEDIEVLDAFIRLKIDPLENADFKVLDLYSPDELSR 611

Qy      670 SNVFDIETMFRYNYAQRVSGRLKDLDDLSTNRNQFVDAFGSLMDDLCAVGQTVNNAVS 729
Db      612 ANVFDLENILREYNSYK---SALYTIKIAKNTPTSYNGINSFLOGGAGTGTGLGVSIS 668

Qy      730 GVATLFSSIVTGFINFKNPFGGLMIIIVIGVLFALYFLTKTKTIYETAPIKMIYP--- 786
Db      669 VTAGALGDIVGGVSWFLKNPFGGLMLILAIIVVWIIIVFVRQKHVLSKPIDMMFPYAT 728

Qy      787 -----EIDKLKREKSEIAPISEEE 807
```


Qy 98 RFOTDHYCPDASDVH-SEGILLIYKONIIPFMFRVRKRKVVTSTVNGIYSDSITNQ 156
Db :
120 RFGKSIDCINHPTKTPVOEGIMVVYKENIVAYTFEVIYHK---DAIFORSYADTTTNY 175
Qy 157 H---TFYKSTEP-WETEKMDTIYOCFNSL-RLNTGNNLLTYVDRDDINMTVFLOPVDGVGT 211
Db :
176 FLGTSVTMAFFPIWELDEOVNRRNCYSAAGRILNGEVVAYHEDSYENYTMVLVEDDYRS 235
Qy 212 PDVKRYGSQBELYLEPCWFNGSVYRRTTVNCELMDMFARSNPPDFRFVTATGDTVMSPF 271
Db :
236 KNSKRYVTTKSRYHKGAWTW-RYTESCNMVCVVVTKARSNTPYEFFVLSSGEVVEISPP 294
Qy 272 WSGDDDHENKMHEKPFPVSVINNVKKVDY-----QNRGTVPGLGKTRIFLDREEVYTLSEWK 326
Db :
295 YNGENSEPFEEDTENFWIR--KNYTKMYTGELAAPKKVVPVPL---MAFLEREDWTIGWEI 349
Qy 327 HLKNWSYCPLTLWKAFYNGIQTEHSYGSHVFANDITASFTTSKEDMKEFNT-----T 378
Db :
350 FPKQNVTCDMKWQTVSRAIRTDNTNTGYHFVSKGLTATFVASKRKI-DYNTTEGKNYNT 408
Qy 379 YHCLNEEIKAEIEKYYA-KVNSTHSKYGDLYKPKTDGGLYLVMQPPLLQNRLLDAKNKLNN 437
Db :
409 FRCVYDFVEBVNRVFBDEYNETHVKDGELEMVRTTGGILVLMQG-----LKAKSLHLNL 462
Qy 438 ETYSRRRRQAESTTDPW-MEMTNG-AGGEYSSEN-----SITVAQVQVAYDNLR 486
Db :
463 EKFAALNNVS VATASPVTTAATENGTTVSRRRKESFDNLDDVVDTSYAQLQFTYDVLK 522
Qy 487 IRINNILEDISKACRQHRAALVMNELSKINPTFSVMISMINYRPVS AKRI GDVISVNCI 546
Db :
523 DYINDALRNIMDAWCROKQRTAEMLKLSKINPSNILSAIVERPVATAKLAGDIAMSECV 582
Qy 547 VVDQTSVSLHKSRLLSASDEK---CESRPPVPFKFNWDSTIYKGLGVANELLTTTYL 603
Db :
583 KVDQSSVKVLKDMRIF-QDGKVVCYSRPLVWFQFINSTKLESGLGENNEIMLGFT 640
Qy 604 ETCQENTFYFOAKTDMVIYKNVBEHLKTVP LSSITLTDTFIALNFTLENVDFKVI ELYT 663
Db :
641 ENCDTNRKIFVVGTVGYEYRDYRPNVTSLEHIDLVDTLIGLDIEPLENTDKVLELYS 700
Qy 664 RDEKRLSNVFDIETPMFREYNYYAQRVSGLRKDLDLSLNQRNFVDAGFSLMDDLGAVGQT 723
Db :
701 KGELRASNVFSLDEIMREYNSQKHIRTLSAKVND--NTPSYLLGLDITFMQGLGVAGKG 757
Qy 724 VNNAVSGVATILFSSIVTGFINFKNPFGGMMLIIWVGVLFAIYFLTCKTKVIETAPIKM 783
Db :
758 IGAIGAAGVAGGVSSVNAVNTGFLTNPGGGFTILLVIGVLAAVYLI P-----TREVEY 810
Qy 784 IYP 786
Db 811 FFP 813

RESULT 16
D48349
glycoprotein B precursor - saimirine herpesvirus 1 (strain MV-5-4-PSL)
C:Species: saimirine herpesvirus 1
C>Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 09-Sep-1994
C:Accession: D48349
R:Eberle, R.; Black, D.
Arch. Virol. 129, 167-182, 1993
A>Title: Sequence analysis of herpes simplex virus gB gene homologs of two platyrrhine
A:Reference number: A48349; MUID:93228440; PMID:8385913
A:Molecule type: DNA
A:Accession: D48349
A:Residues: 1-920 <EBE>
A>Note: sequence extracted from NCBI backbone (NCBIN:129066, NCBI:P129068)
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: Glycoprotein; transmembrane protein
F:1-35/Domain: signal sequence #status predicted <SIG>
F:36-920/Product: glycoprotein B #status predicted <MAT>
F:576-592/Region: hydrophobic

A;Title: Nucleotide sequence and characterization of the Marek's disease virus homologue
A;Reference number: A32402; MUID:89293086; PMID:2544666

A;Accession: A32402
A;Molecule type: DNA

A;Residues: 1-865 <ROS>

A;Cross-references: GB:D13713; EMBL:D00506; NID:g221836; PIDN:BAA02866.1; PID:g221837

A;Accession: B32402

A;Molecule type: protein

A;Residues: 250-271;304-330 <ROS2>

C;Superfamily: herpesvirus glycoprotein B

C;Keywords: glycoprotein; transmembrane protein

F;1-21/Domain: signal sequence #status predicted <SIG>

F;22-865/Product: glycoprotein B #status predicted <GPB>

F;709-728/Domain: transmembrane #status predicted <TN1>

F;732-752/Domain: transmembrane #status predicted <TN2>

F;27,184,332,364,406,425,631/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 19.3%; Score 876.5; DB 1; Length 865;

Best Local Similarity 27.0%; Pred. No. 3.4e-44;

Matches 231; Conservative 164; Mismatches 386; Indels 75; Gaps 16;

QY 51 YGTHDSHGGERNNRDSSEQNKIYGPSPTFPYRVCSASGVGVFRFQTDHVCPSDASD 110

DB 17 YGTHNSP-----STQNVTSREVSVSSVQSEESTFYLCPPVPVGVSTVIRLEPPRKCPEPRK 71

QY 111 MVH-SEGILLIYKQNIIPFMRVRYKRVKVVTTSTVNGIYSDSIHQHTFYKSIPEWE-T 168

DB 72 ATEMGEGIALPKENISPKFKVLYKNIQITTTGTGYQIINRYTDRTPVSEIET 131

QY 169 EKMDTIYQCFNSLRNLNTGNNLITYVDRDDINNTVFLQPDVGVTPVVKYSGOPELYLEFG 228

DB 132 DLIDGKGRCSKARYLRNVVYVEAFDRDAGERQVLLKPSKENTPESRAWHTTNITYT--- 188

QY 229 WFMGS---YRRRTVNCCLMDMFARNPPDFEVATGDTVMSFPW--SGEDDHENKMH 283

DB 189 -VWGSPWIRVGTGTVNCVIEEMDARSVPFYSYFAMANGDIANISPFYGLSPPEAAAEPMG 247

QY 284 EKPWFVSVNNYKVDYQNRGVTVPLGKTRIFDRBEYTLSEKHLKNSYCPCLTLWKAFY 343

DB 248 YQDNFKQLDSYFSDMLDKRKASLPVKENFLITSHFTVGWDWPKTRVCSMTKKEVT 307

QY 344 NGIQTEHSGSYHFVANDITASTTTKEDMKENTTY-----HCLNEEKABIEKKY 394

DB 308 EMLRATVNGRYRFMARLEISATFIS-----NTTEFDNRILIGQCIRKREAAAEQIF 359

QY 395 -AKVNSTHSGYGLKFKYKTDGGLYLWQPLIQ-----NRLLDKAKLN 436

DB 360 RTKYNDSHVKVGHVQYFLALGQFVAYQPVLSKSLAHMYLRELMDNRDTEMLDLVNNKH 419

QY 437 -----NETYSRRSRQAESETDPMEMTNGAGGEYSSENSITVAQVQVAYDNLRIINN 491

DB 420 AIYKKNATSLRLRDIRNAPRKITLDDTTA---IKSTSSVQFAMQLFYDHIOTHIND 476

QY 492 ILEDLSKAWCREQRAALVWNLKINPTSVMSIYNRPVSAKRIGDVISVNCIVVDQT 551

DB 477 MFSRIATANCELQNLRELWLWHEGINKINPSATASATLGRRAAKLGDVAASVSSCTAIDAE 536

QY 552 SVSLHSLRLLSASDEKCSRPPVTFKFNDSITYKQLGVNNEILLITTYLETQENTE 611

DB 537 SVTLQNSMRVIT--STNYTCYSRLVLFVSYGENOGNIQQLGENNELLPLEAVEPCSAHR 595

QY 612 VYFOAKTDMYIKVNEHLKTVPLSSITTLDTLIALNFTLLENVDKVIETYLRDSKRLSN 671

DB 596 RYFLPGSGVALPENYVFKVDAADQIASTFVELNLTLEDEILPISVYTKELRDVG 655

QY 672 VFDIETMFREYNYYAQRVSGRLKDLDSLSTNRNQFVDAGFGLMDDLGAVQTVNNAVSGV 731

DB 656 VLDYAEVARRNQLHELKFDYDINK-VIEVDTN-YAFMNGLAELFNGMGQVQGAIGVVGA 713

QY 732 ATLFSSIVTGFINFKNPPGGMMLIIVGLVFAFYFLTKTKTIYETAPIKMIYPIDKL 791

DB 714 AGAIVSTISGVSFAFMSNPFAGALIGLIIAGLVAFLAYRYVNNKLKSNPMKALYPMTEV 773

QY 792 KERECKSEIAPISBEELER-----IVLAMHIHQONGHMETTKRDKPKDSILT 838

DB 774 LKAQATRELHGEESDLDLRTSIDERKLEAREMIKYMALVSAERHEKRLKRRGTTAV 833

QY 839 RAQNLKRRSGYSNLK 854

DB 834 LSDHLAKMRIRKSNPK 849

RESULT 18

A56602

glycoprotein B homolog precursor - feline herpesvirus 1

C;Species: feline herpesvirus 1

C;Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 26-Aug-1999

C;Accession: A56602

R;Maeda, K.; Horimoto, T.; Norimine, J.; Kawaguchi, Y.; Tomonaga, K.; Niikura, M.; Kai,

Arch. Virol. 127, 387-397, 1992

A;Title: Identification and nucleotide sequence of a gene in feline herpesvirus type 1

A;Reference number: A56602; MUID:93090104; PMID:1333759

A;Accession: A56602

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-948 <MAE>

A;Cross-references: GB:S49775; NID:g261094; PIDN:AAB24381.1; PID:g261095

A;Note: sequence inconsistent with nucleotide translation

A;Note: sequence extracted from NCBI backbone (NCBIN:120002, NCBI:120003)

C;Superfamily: herpesvirus glycoprotein B

C;Keywords: glycoprotein

Query Match 19.0%; Score 862; DB 2; Length 948;

Best Local Similarity 27.3%; Pred. No. 2.8e-43;

Matches 244; Conservative 165; Mismatches 387; Indels 98; Gaps 24;

QY 24 SLSTAETGVTSPPTATWSTESPLTGHYGHDSHGGERNNRDSERQNKIYGPSPTF 83

DB 85 TVATPEVGGT-PPKPTTPTDPTD-----SDMREALRASQIEANGPSTF 125

QY 84 PYRVCSASGVGVFRFOTDHFVCPDAS-DMVHSEGLIYKQNIIPFMRVRYKRVKVVTT 142

DB 126 --YWCPPPSGTVVRLPEPPRACPDYKLGKNTFEGIAVIFKENIAPYKANIYKNIIMT 183

QY 143 TVYNGIYSDSIHQHTFYKSIPEWE-TEKMDTIYQCFNSLRNLNTGNNLITYVDRDDINMT 201

DB 184 TVWSSSVAVTNYTDRVPVKQVEITDLIDRRGWCLSKADYVRNNYQTFATDRDEDRE 243

QY 202 VFLQPDVGVTPDVVKYSGOPELYLEPCGFWGYSRRRTVNCCLMDMFARNPPDFFTVA 261

DB 244 LPLKPSKFNTPQSRGWHITTNETYTKIG-AAGFHHSHTSVNCIVEEDARSVYPYDSFAIS 302

QY 262 TGDVTMSFPWSGED---DHENKMKPMPFVSVINNYKVDYQNRGVTVPLGKTRIFLDR 317

DB 303 TGDVHMSFPFLGRDGAHVHTSYSSDR---FQIIEGYPIDDLTRQLQGPVGRNFLET 359

QY 318 BEYTLSEKHLKNSYCPCLTLWKAFYNGIQTEHSGSYHFVANDITASFTTSKEDMKFENT 377

DB 360 PHVTIVANNWTPKCGRCVCTLAKWEIDEMLRDEYQGSYRFTVKTTISATFISNTSQF-EINR 418

QY 378 TY--HCLNEEKABIEKKY-AKVNSTHSGYGLKFKYKTDGGLYLWQPLI----- 424

DB 419 IRLGDCATKEAAEAIDRIYKSKYKTHIQGTLETYLARGGFLIAFRPMISNELAKLYIN 478

QY 425 ---ONRLLDKAKLN--NETYSR-----RSRQAESETDPMEMTNGAGGEY 467

DB 479 ELASRNTVDLSALNPSGETVQTRGVSFNSQNHRSRR---STIEGIEVNNAS--LL 533

QY 468 SSENSITVAQVQVAYDNLRIINNILEDLSKAWCREQRAALVWNLKINPTSVMSMIY 527

DB 534 KTTSSVEFAMIQPAYDIQAHVNEMLSRITATWCTLQNRHVLWTETLKLNPGGVSMAL 593

QY 528 NRPVSARKRIGDVISVNCIVVDQTSVSLHSLRLLSASDEKCFRPPVTFKPMNDSTIYK 587

DB 594 ERRVSARLLGDAVAVTQCVNISSGHVYIQNSMR-VTGSSTTCYSRPLVSLFRALNDSEYIE 652

Query Match	18.6%	Score	845.5	DB 1	Length	903			
Best Local Similarity	27.7%	Pred. No.	2.5e-42						
Matches	245	Conservative	179	Mismatches	359	Indels	101	Gaps	28
Qy	28	AETGVTSPNATATWSTESP--LTGHYGHDSHGE-RCNNENRROSECKNIYGSPTFP	84						
Db	61	APTGDTPKKNKKPNPPPPAGDNATVAAGHATLRHLRDIKAENTDANFY-----	113						
Qy	85	YRVCSAGVGDFRFTQDHDVCPDASD-MVHSEGILLIYKQNIIPFMFRVRYKRVVTTST	143						
Db	114	--VCPPTGATVQFEQPRRCPTREPGQNYTEGIAVVFKENIAPYKPKATMYKDVTSQ	171						
Qy	144	V-----YNGIYSDSITNOHTFYKSTPEW--TEKMDTIYQFNSLURLTGNLLTYV	193						
Db	172	VWFHRYYSQFMGIFED-----RAPVPFEVIDKINAGVCRSTAKYVRNNLETTAF	222						
Qy	194	DRDDINMTVFQPDVGV-----TPDKRYGSGPELYLEPGFWGYSYRRTTVCNCELM	245						
Db	223	HRDDHETDMELKPANAAATRTSRGWHHTDLKYNPSVEAF-----HRVGTTVNCIIVE	273						
Qy	246	DMFARSNPPDFFTVATGDTVEMGPFWS-GBDDH-ENKQHEKPMFVSVINNY-KVVDYQN	302						
Db	274	EVDARSVVPYDFVLATGDFVYMSPFYRGSGSHEHTSYAADRFKQVDGFYARDLTTKA	333						
Qy	303	RGTVPLGKTRIFLDREBYTLWSKHLKMSYCPCLTLWKAFYNGIQTEHSGSYGHFVANDIT	362						
Db	334	RATAP--TTRNLLTTPKTFVAMDVWPKRPSVCTMTKQOEVDMLRSEYGGSRFSSDAIS	391						
Qy	363	ASFTT-----SKEDMKFEFTTYHCLNEEBIKAEITEKYA-KVNSTHSKYGDLYKFKTDG	414						
Db	392	TTFTTNLTPEPLSRVDLGD-----CTGKDARDAMDRIFARYNATHIKVGQPOYYLANG	445						
Qy	415	GLYLWVOPLIONRLLDKKNLNNETYSRRSRROAESTTDPNMENTGNGAGGEYSSENSIT	474						
Db	446	GFLIAYQPLLSNTLAE---LYVREHLREQSRKPNPTPPPPGASANASVERIKTTSSIE	501						
Qy	475	VAQVOYAYDNLRIBINNILEDLSKAWCREQHRAAALVNWELSKINPTSVMSMIYRNPVSAK	534						
Db	502	FARLOFTYNIHQHVNDMLGRVAJAWCELONHLETLWNEARKLPNPAIASATVGRVRSVA	561						
Qy	535	RIGDIVGSNCIVVDQTSVLHSLKRLLSASDEKCFSRPPVTPFKFMDSTIYKQGLGVNN	594						
Db	562	MLGDVMAVSTCPVAADNVIQNSWR-ISSRPGACYSRPLVSFRYEDOGPIVEQOLGNN	620						
Qy	595	EILTTTYLETQEBENTYYFOAKTDMYIKNYEHLKTVPPLSSITLDTLFIALNFTLLENV	654						
Db	621	ELRLTRDAIBPCTVGHRRYPTFGGVYVFEYAYSHQLSRADITTVSTTFIDNLITMLEDH	680						
Qy	655	DFKVIELYLRDEKLSNVFIETWFRVNYIAQVSGULRKDLLDLSNTRNOF--VDAPGS	712						
Db	681	EFVPLEYVTRHEIKDGLDYTEQRRNQLHDLRFADI-DTVIHADANAAMFAGLGAFFE	739						
Qy	713	LMDDLG-AVGOTVNVAVSGVATLTFSSITVGFINKPFPFGMLMIIVVIGVLFAIYPLTK	771						
Db	740	GMGDLGRAVKVMGIVGGVV-----SAVSGVSSFMSNPPFGALAVGLLVLAGLAAAFPAFR	795						
Qy	772	KTKIYETAPIKMIYPEIDK-LK-----EREKSEIAPISE-EELERIVLAMHIH	818						
Db	796	YVMRLQSNPMKALYPLTTTKELKNTPDASGEGEEGGDFDEAKLAAREMIRYMALVSAM	855						
Qy	819	QONSHMETKTRKDPKDSILT-RAQNML---KRRSGYSNLKNAE	857						
Db	856	ERTEH---KAKKGTSTALLSAKVDTMVMRKERNNTYQVQPKD	895						

RESULT 21
VGBEK1
glycoprotein B precursor - human herpesvirus 1 (strain KOS)
C/Species: human herpesvirus 1
C/Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999
C/Accession: A03751
R/Strik, D.J.; Fox, B.A.; DeLuca, N.A.; Person, S.
Virology 133, 301-314, 1984

A:Title: Nucleotide sequence specifying the glycoprotein gene, gB, of herpes simplex vir
A:Reference number: A03751; MUID:84174058; PMID:6324454
A:Accession: A03751
A:Molecule type: DNA
A:Residues: 1-903 <BZ>
A:Cross-references: GB:K01760; NID:g330082; PIDN:AAA5774.1; PID:g330083
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein; transmembrane protein
F:1-30/Domain: signal sequence #status predicted <SIG>
F:31-903/Product: glycoprotein B #status predicted <MAT>
F:726-746/Domain: transmembrane #status predicted <TM1>
F:751-771/Domain: transmembrane #status predicted <TM2>
F:774-794/Domain: transmembrane #status predicted <TM3>
F:76:86,140,254,397,429,477,488,673,818,887/Binding site: carbohydrate (Asn)
F:115-572,132-528,206-270,363-411,595-632/Disulfide bonds: #status predicted
F:115-572,132-528,206-270,363-411,595-632/Disulfide bonds: #status predicted (covalent)

Query Match	18.6%	Score 845.5	DB 1	Length 903
Best Local Similarity	27.6%	Pred. No. 2.5e-42		
Matches	245	Conservative 181	Mismatches 360	Indels 101
Gaps	28			
Qy	24	SLSLAETGVTSPPNTATWSTESP--LTGHVGHTHDSSHGE--RGNNENRDSEONKNITGSP	80	
Db	57	ALGAAPTGDPKPKKNKKPKNPTTPRAGDNATVAAGHATLREHLRDKAENTDANFY---	113	
Qy	81	STFPYRVCSASGVGDVRFQFDTHDVCPSADP--MYHSEGLIILYKONIIPIPMFRVRYKRVV	139	
Db	114	-----VCPPTGATVVQFPQPRRCPTREPQONVTEGIAVVFKENIAPYKFKATMYKDV	167	
Qy	140	TTSTV-----YNGIYSDSIINQHTFYKSIPEWE--TEKMDTIYQCFNSLRNTGGNL	189	
Db	168	TVSQVMFGHRYSGFMGIFED-----RAPVFEEVDIKNAKGVCSTAKYVRNNLE	218	
Qy	190	LTVVDRODINMTVFLQPDGV-----TPDKRYGSOPELYLEPGWFGWSYRRRTTVN	241	
Db	219	TTAFHRDDHETMELKPANAATSTRGWHTTDLKYNPSRVEAF-----HRYGTTVN	269	
Qy	242	CELMDMPARPNPDPFVATGTOTVEMSPWIS--GEDDH--ENKMKHPWFVSVINNY--KV	298	
Db	270	CIVEEVDARSVYPDEFVLATGDFVVMSPFYGYREGSHTHTTYAADRFPQVDGFYARDL	329	
Qy	299	DYQNRGTVPUGKTRIFLDREBYTLSEKHLKNMSYCPLTLKPAFYNGIOTEHSGSYHFVA	358	
Db	330	TTKARATAP--TTRNLLTTPKFIVANDWVPKPSVCTMTKQWDEMLSEYGCSPRFS	387	
Qy	359	NDITASFTT-----SKEDMKFNTYTHCLNBEIKAEIEKYA--KVNSTHSKYGDLYKF	410	
Db	388	DAISTTFTNTLNEYPLSRVDLGD-----CIGKADAMDRIEFARRYNAITHIKVGQPOY	441	
Qy	411	KTGGLYLWMOPLIQNRLLDAKKNLANNETYSRRSRQAESTTDPMMEMTCNGAGGEYSSE	470	
Db	442	LANGGELIAYQPLLSNTLAE---LYVRHLEQSRKPNPTPPPGSANAASVERIKTT	497	
Qy	471	NSITVAQVOYAYDNRIRINNILEDLSKAWCRQHRAALVWNELSKINTPTSVMSMITYNRP	530	
Db	498	SSIEFARLQPTYNHIIQRHVNDMLGRVAIACWELQNHLETLWNEARKLPNPAIASVTVGR	557	
Qy	531	VSAKRGDVLUSNCIVDQTSVLSKSLRLLSASDEKCFSRPPTVTFKPMNDSTIYKQQL	590	
Db	558	VSARMGLDGMVASTCVPVADNVIQVNSMR--ISSRFGACYSPLVSRFYREDQGPLVEGOL	616	
Qy	591	GVNNEILLTTTYLETQCENTEYVFOAKTDWYIKNYEHLKTVPLSSITTLDTFIALNFTL	650	
Db	617	GENNELRLTRDALEPCTVGHRRYFTPGGVYVFEYAYSHQLSRADITTVSTVFIDLNI	676	
Qy	651	LENVDPKVIELYTRDEKRLSNVFDIETMPREYNYYAQRVSGRLRKDLLDLSGRNQF--VD	708	
Db	677	LEGHEFVPLEVYTRHEIKDSGLLDYTEVQRNQLHDLRPAADI--DTVIHADANAAMPAGLG	735	
Qy	709	AFGSLMDDLGL--AVGQTVNVNASGVATLFSIIVTGTFINFIKNBPFGGMLMIIVIGVLFAIY	767	
Db	736	AFPEGMGDLGRAVGKVMGLVGWV-----SAVSGVSSPMSNPFGAUAVGLVLGAAAF	791	
Qy	768	FLTKTKYIVETAPIKMIYPEIDK-LK-----EREGKSEIAPISE-BELERIVLA	814	

```
Db      792  FAFYVMRLQSNPKALYPLTTKELKNPTNPDPASGESEGGDFDEAKLAAREMIRYMAL 851
QY      815  MHIHQSNHMETKTRKDPKDSILT-RAQNML-----RKRGSGYNLKNAE 857
Db      852  VSAMERTEH---KAKKGTSLRLSAAKVTDMVMKRKRNTNTQVFNKD 895

RESULT 22
VGBEW7
glycoprotein B precursor - human herpesvirus 1 (strain 17)
C:Species: human herpesvirus 1
C:Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 16-Jun-2000
C:Accession: I30084
R:McGeoch, D.J.; Dalrymple, M.A.; Davison, A.J.; Dolan, A.; Frame, M.C.; McNab, D.; Petr
J. Gen. Virol. 69, 1531-1574, 1988
A:Title: The complete DNA sequence of the long unique region in the genome of herpes sim
A:Reference number: A30083; MUID:88274327; PMID:2839594
A:Accession: I30084
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-904 <MCG>
A:Cross-references: GB:X14112; GB:D00317; GB:D00374; GB:S40593; NID:g1944536; PIDN:CAA32
C:Genetics:
A:Gene: UL27
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: Glycoprotein; transmembrane protein
F:1-30/Domain: signal sequence #status predicted <SIG>
F:31-904/Product: glycoprotein B #status predicted <MAT>
F:727-746/Domain: transmembrane #status predicted <TM1>
F:752-771/Domain: transmembrane #status predicted <TM2>
F:774-795/Domain: transmembrane #status predicted <TM3>
F:87,141,255,398,430,478,489,674,819,888/Binding site: carbohydrate (Asn) (covalent) #sq
F:116-573,133-529,207-271,364-412,596-633/Disulfide bonds: #status predicted

Query Match      18.5%; Score 841; DB 1; Length 904;
Best Local Similarity 27.7%; Pred. No. 4.7e-42;
Matches 242; Conservative 179; Mismatches 345; Indels 108; Gaps 28;

QY      35  PPNTATWSTSPLTGCHYTHDSSHGE-RGNENNRDSEQKNYIGSPSTPYRVCSASGV 93
Db      80  PPRPA-----GDNATVAAGHATLRHLRDIKAENTDANFY-----VCPPTG 121
QY      94  GDVRFQTDHVCDDASD-MVHSEGLIYKQNIIPFMRVRYKRVKVTSTV----- 144
Db      122  ATVVQFQPRCPTRPQGNQTEGIAVFNKENTAPYKFKATMYKDVTVSQVWFHRYSQ 181
QY      145  YNGIYSDSITNQHTFYKSIQBPWE--TEKMDTIYQCFNSRLNTGGNLLTYVDRDDINMTV 202
Db      182  FMGIFED-----RAPVPFEVIDKINAKGVCSTAKYVNNLETTAFHRDDHETDM 232
QY      203  FLOPVDGV-----TPDVRYGSGQPELYLEPGWFWGYSYRRRTTNCELMDMFARSNPP 254
Db      233  ELKPANAATRTSRGWHTTDLKYNPSVEAF-----HRYGTTVNCIVEEVDARSVYP 283
QY      255  FDFPVTATGTVEKSPWS-GEEDH-ENKMKHEKPFVSVINNY-KVVDYQNRGTVPVGKT 311
Db      284  YDEFLVATGDFVYNSPPYRGESHTHTSYAADRFPQVDFGFIARDITTKARATAP--TT 341
QY      312  RIFLDREYILTSWEKHLKNMNSCYPLTLWKAFYNGIQTEHSGSYHFVANDITASFTT--- 367
Db      342  RNLLTTPKFTVAMDVPKRSVCTMTKQVEDEMLRSEYGGSPFSSDAISTFTTNLTE 401
QY      368  ---SKDMKFNTHYCLNNEIEKAEIEKKYA-KVNSTHSKYGDLKYFKTDGGLYLVNQPL 423
Db      402  YPLSRVLDGD-----CIGKDARDAMDRIFARRYNATHIKVGPQYQVYLANGGELIAYQPL 455
QY      424  IQNELLDKAKLNNETYSRRSRQAESTTDPMMEMTCNGAGGEVSSSENSITVAQVOYAYD 483
Db      456  LSNLTAB-----LYVREHLRQSRKPPNPTPPPGASANASVERIKTSSIEFARLQPTYN 511
QY      484  NLRIRINILEDLSKAWCRQHRAALVNMELSKINPTSVMSMTIYNRPVSARKIGDVISVS 543
```

```
Db      512  HIQRHVNDMLGRVAIAJAWCEQLQNHELTMNWEARKLNPNALIASATVGRVRVARMGLDGVAVS 571
QY      544  NCIVVDQTSVLSHLKSLRLSASDEKCFSPPTVFTEKEMNDSTIYKGGOLGVNNEILLTTTTL 603
Db      572  TCVFVAADNVIVQNSMR-ISSRPGACYSRPLVSFRYEDQGLVEGQNNELRLTRDAI 630
QY      604  ETCQENTETTYFOAKTDMYIKNYEHLKTVPLSSITLDTFFIALNFTLLNVDPKVIELYT 663
Db      631  EPTCVGHRRYFTFGGYYVFEYAYSHQSLRADITTVSTVFIDNLNITMEDHEFPVLEVYT 690
QY      664  RDEKRLSNVFDIETMFREYNYAQRVSGRLKOLLDLSTNRNQF--VDAFGLMDDLG-AV 720
Db      691  RHEIKDGLDVTQVRRNQLHDLRPADI-DTVIHADANAAMPAGLGAFFEGMGDLGRAV 749
QY      721  GQTVNVNAGSVATLPSSIVTGFINFKNPFGGMLMIIIVIGLVFAIYFLTKTKTIVETAP 780
Db      750  GKVMGIVGGV-----SAVSGVSSFPNSNPFALAVGLLVLAGLAAFAFFRYVMRQSNP 805
QY      781  IKMIYPEIDK-LK-----EREGKSEIAPISE-BELERIVLAMHIHQONSHMETK 827
Db      806  MKALYPLTTKELKNPTNPDPASGESEGGDFDEAKLAAREMIRYMALVSAMERTEH--K 862
QY      828  TRKDPKDSILT-RAQNML---RKRGSGYNLKNAE 857
Db      863  AKKKGTSALLSAKVTDMMVMKRKRNTNTQVFNKD 896

RESULT 23
VGBEB2
glycoprotein B precursor - human herpesvirus 2 (strain 333)
C:Species: human herpesvirus 2
C:Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 16-Jul-1999
C:Accession: A26790; A45928
R:Stuve, L.L.; Brown-Shimer, S.; Pachl, C.; Najarian, R.; Dina, D.; Burke, R.L.
J. Virol. 61, 326-335, 1987
A:Title: Structure and expression of the herpes simplex virus type 2 glycoprotein gB gene
A:Reference number: A26790; MUID:87112925; PMID:3027364
A:Accession: A26790
A:Molecule type: DNA
A:Residues: 1-904 <STU>
A:Cross-references: GB:M15118; NID:g330256; PIDN:AAA45837.1; PID:g330257
R:Zwaagstra, J.C.; Leung, W.C.
Can. J. Microbiol. 33, 879-887, 1987
A:Title: The nucleotide sequence of herpes simplex virus type 2 (333) glycoprotein gB2 an
A:Reference number: A45928; MUID:88079667; PMID:2446730
A:Accession: A45928
A:Molecule type: DNA
A:Residues: 1-34, 'AMPTV', 42-307, 'T', 309-481, 'R', 483-609, 'M', 611-664, 'R', 666-904 <ZMA>
A:Cross-references: GB:M24771; NID:g341245; PIDN:AAA60540.1; PID:g623400
A:Note: the authors translated the codon ATG for residue 610 as Ile
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein; transmembrane protein
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-904/Product: glycoprotein B #status predicted <MAT>
F:724-744/Domain: transmembrane #status predicted <TM1>
F:749-769/Domain: transmembrane #status predicted <TM2>
F:772-792/Domain: transmembrane #status predicted <TM3>
F:82,136,250,373,425,473,486,671/Binding site: carbohydrate (Asn) (covalent) #status predi
F:111-570,128-526,202-266,359-407,593-630/Disulfide bonds: #status predicted

Query Match      18.5%; Score 839; DB 1; Length 904;
Best Local Similarity 28.3%; Pred. No. 6.2e-42;
Matches 246; Conservative 164; Mismatches 367; Indels 92; Gaps 27;

QY      33  TGPPNTATWSTSPLTGCHYTHDSSHGERGNENNRDSEQKNYIGSPSTPYRVCSASG 92
Db      76  TPEPD---ANATVAAGHATLR-AHLREIKVENADAQ-----FYVCPPT 115
QY      93  VGDVRFQTDHVCDDASD-MVHSEGLIYKQNIIPFMRVRYKRVKVTSTV----- 144
Db      116  GATVWQFQPRCPTRPQGNQTEGIAVFNKENTAPYKFKATMYKDVTVSQVWFHRYQS 175
QY      145  YNGIYSDSITNQHTFYKSIQBPWE--TEKMDTIYQCFNSRLNTGGNLLTYVDRDDINMT 201
```

```
Db 176 QFMGIFED-----RAPVPEEVIDKINAKGVCRSTAKYVRNNMETTAFHRDDHETD 226
Qy 202 VELQPDVGTPTDKRYGSOPELVEGFWGSRRTTUNCCLMDMFARSNPPDFVFTA 261
Db 227 MELPKAK-VATRTSRGWHTTDLKYNPSRVEAFHRYGTTVNCIVEEDARSVPYDFEFLA 285
Qy 262 TGDVTMSPFWS-GEDDH-ENKMHKPMFVSVINNY-KVVDYQNRGTVPGLKTRIFLDRE 318
Db 286 TGDVYMSPPFYGYREGSHTEHTSYAADRFKQVDGFYARDLTTKARATSP--TTRNLLTTP 343
Qy 319 EYTLSEKHLKMSYCPFLWKAFYNGIQTEHSGSYHFVANDITASFTT-----SKED 371
Db 344 KETVAMDWPVKRPAPVCTMTKQVEDEMLRAEYGSFRFSSDAISITFTTNLTQYSLSRVD 403
Qy 372 MKEFNNTYHCLNEIEKAEIEKKYA-KVNSTHSGYGLKYFKTGDGLYLWQPLIONRLLD 430
Db 404 LGD-----CIGDAREALDRMFARKYNATHIKVGOPYLYLATGGFLIAYQPLLSNTLAE 457
Qy 431 AKNKLNNETYSR-RSRROAESTTDDPMEM-TGNAGGEYSSENSITVAQVQYAYDNLRIR 488
Db 458 ----LVVREYMRQDRKPRNATPAPLREAPSANASVERIKTSSIEFARLQTYNHIQRH 513
Qy 489 INNILEDKAMCREQHRALVWNLKINPTSVMSIYNRPVSAKRIGDVISVNCIVV 548
Db 514 VNDMLGRIAVAWCELQNHLETLWNEARKLNPNAIASATVGRRVSAARMGLGDVMAVSTCPVP 573
Qy 549 DQTSVSLHSLRLLSASDEKCFSPRPVTFKPMNDSTIYKQGLGVNNEILLTITTYLETQOE 608
Db 574 APDNVIVQNSMR-VSSRPGTCYSRPLVSFRYEDOGPLIEGQLGNNELTRDALEPCTV 632
Qy 609 NTEYFQAKTDMYIKNYEHLKTVPLSSITLDTFIALNFTLLENVDPKVIELYTRDEKR 668
Db 633 GHRGYFIFGGYVYFEYAYSHQLSRADVTTVSTFIDLNTMLEDHFEFVPLEVYTRHEIK 692
Qy 669 LSNVFIETMFREYNYAQRVGLRKLDDLSTNRNQFVD--AFGSLMDDLG-AVGQTVV 725
Db 693 DSGLLDYTEVQRRNQLHDLRFADI-DTVIRADANAAMFAGLCAFFEGMGDLGRAVGKVM 751
Qy 726 NAVSGVATLFPSSIVTGFINKPFGGMLMIIVIGVLFALYFLTKTKIYETAIKMIY 785
Db 752 GVUGGVV----SAVGSVSSFMSPFGALAVGLLVLAGLVAAFFAFRYVLQLOQNPMAKY 807
Qy 786 PEIDKLKEREKSEIAPISSE-----ELERIVLAMIHIQOONSHMETTKRDP 832
Db 808 PLTTKELKTSDFGGVGEEGEGGDFEAKLAAREMIRYALVAMSERTEHKARKKG 867
Qy 833 KDSIL-TRAQNM-LRKR--SGYSNLKNAE 857
Db 868 TSALLSSKVTNMVLRKRKARYSPLHNE 896
```

RESULT 24

VBGEK2

glycoprotein B precursor - human herpesvirus 2 (strain HG52)

C.Species: human herpesvirus 2

C.Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 16-Jul-1999

C.Accession: A25611

R:Bzik, D.J.; Debroy, C.; Fox, B.A.; Pederson, N.E.; Person, S.

Virology 155, 322-333, 1986

A.Title: The nucleotide sequence of the gB glycoprotein gene of HSV-2 and comparison with

A.Reference number: A25611; MUID:87071654; PMID:3024391

A.Accession: A25611

A.Molecule type: DNA

A.Residues: 1-904 <BZI>

A.Cross-references: GB:M14923; NID:G330254; PIDN:AAA66440.1; PID:G330255

R.Norais, N.; Tang, D.; Kaur, S.; Chamberlain, S.H.; Maslars, F.R.; Burke, R.L.; Marcus,

J. Virol. 70, 7379-7387, 1996

A.Title: Disulfide bonds of herpes simplex virus type 2 glycoprotein gB.

A.Reference number: A58366; MUID:97048015; PMID:8892856

A.Contents: annotation; tryptic peptide disulfide bond assignments

C.Superfamily: herpesvirus glycoprotein B

C.Keywords: glycoprotein; transmembrane protein

F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-904/Product: glycoprotein B #status predicted <MAT>
F:724-744/Domain: transmembrane #status predicted <TM1>
F:749-769/Domain: transmembrane #status predicted <TM2>
F:772-792/Domain: transmembrane #status predicted <TM3>
F:82,136,250,393,425,473,486,671/Binding site: carboxylate (Asn) (covalent) #status pr:
F:111-570,128-526,202-266,359-407,593-630/Disulfide bonds: #status experimental

Query Match 18.4%; Score 837; DB 1; Length 904;

Best Local Similarity 28.1%; Pred. No. 8.1e-42;

Matches 244; Conservative 167; Mismatches 366; Indels 92; Gaps 27;

Qy 33 TSPNTATWSTESPLTGHYTHDSSHGGRGNENRDSSEONKNIYGSPTSPPYRVCSASG 92

Db 76 TTPPD-----ANATVAAGHATVR--AHLREIKVENADAQ-----FYVCPPT 115

Qy 93 GVDVRFOTDHYVCPDASP-MVHSGILLIYKONIIPMFVRVRYKRVKVVTTTV----- 144

Db 116 GATVVQFEQPRRCPTREPQNYTEGIAVVFKEPIAIFYKATMYKDVTVSQVWFGHRS 175

Qy 145 -YNGIYSDSIINQHTFYKSIETPWE--TEKMDTIYQCNSLRNLNTGGNLLTYVDRDDINMT 201

Db 176 QFMGIFED-----RAPVPEEVIDKINAKGVCRSTAKYVRNNMETTAFHRDDHETD 226

Qy 202 VFLQPDVGTPTDKRYGSOPELVEGFWGSRRTTUNCCLMDMFARSNPPDFVFTA 261

Db 227 MELPKAK-VATRTSRGWHTTDLKYNPSRVEAFHRYGTTVNCIVEEDARSVPYDFEFLA 285

Qy 262 TGDVTMSPFWS-GEDDH-ENKMHKPMFVSVINNY-KVVDYQNRGTVPGLKTRIFLDRE 318

Db 286 TGDVYMSPPFYGYREGSHTEHTTYAADRFKQVDGFYARDLTTKARATSP--TTRNLLTTP 343

Qy 319 EYTLSEKHLKMSYCPFLWKAFYNGIQTEHSGSYHFVANDITASFTT-----SKED 371

Db 344 KETVAMDWPVKRPAPVCTMTKQVEDEMLRAEYGSFRFSSDAISITFTTNLTQYSLSRVD 403

Qy 372 MKEFNNTYHCLNEIEKAEIEKKYA-KVNSTHSGYGLKYFKTGDGLYLWQPLIONRLLD 430

Db 404 LGD-----CIGDAREALDRMFARKYNATHIKVGOPYLYLATGGFLIAYQPLLSNTLAE 457

Qy 431 AKNKLNNETYSR-RSRROAESTTDDPMEM-TGNAGGEYSSENSITVAQVQYAYDNLRIR 488

Db 458 ----LVVREYMRQDRKPRNATPAPLREAPSANASVERIKTSSIEFARLQTYNHIQRH 513

Qy 489 INNILEDKAMCREQHRALVWNLKINPTSVMSIYNRPVSAKRIGDVISVNCIVV 548

Db 514 VNDMLGRIAVAWCELQNHLETLWNEARKLNPNAIASATVGRRVSAARMGLGDVMAVSTCPVP 573

Qy 549 DQTSVSLHSLRLLSASDEKCFSPRPVTFKPMNDSTIYKQGLGVNNEILLTITTYLETQOE 608

Db 574 APDNVIVQNSMR-VSSRPGTCYSRPLVSFRYEDOGPLIEGQLGNNELTRDALEPCTV 632

Qy 609 NTEYFQAKTDMYIKNYEHLKTVPLSSITLDTFIALNFTLLENVDPKVIELYTRDEKR 668

Db 633 GHRGYFIFGGYVYFEYAYSHQLSRADVTTVSTFIDLNTMLEDHFEFVPLEVYTRHEIK 692

Qy 669 LSNVFIETMFREYNYAQRVGLRKLDDLSTNRNQFVD--AFGSLMDDLG-AVGQTVV 725

Db 693 DSGLLDYTEVQRRNQLHDLRFADI-DTVIRADANAAMFAGLCAFFEGMGDLGRAVGKVM 751

Qy 726 NAVSGVATLFPSSIVTGFINKPFGGMLMIIVIGVLFALYFLTKTKIYETAIKMIY 785

Db 752 GVUGGVV----SAVGSVSSFMSPFGALAVGLLVLAGLVAAFFAFRYVLQLOQNPMAKY 807

Qy 786 PEIDKLKEREKSEIAPISSE-----ELERIVLAMIHIQOONSHMETTKRDP 832

Db 808 PLTTKELKTSDFGGVGEEGEGGDFEAKLAAREMIRYALVAMSERTEHKARKKG 867

Qy 833 KDSIL-TRAQNM-LRKR--SGYSNLKNAE 857

Db 868 TSALLSSKVTNMVLRKRKARYSPLHNE 896

```
RESULT 25
VGBECS
glycoprotein B precursor - infectious laryngotracheitis virus (strain SA-2)
C;Species: infectious laryngotracheitis virus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Sep-1994
C;Accession: A40567
R;Kongsuwan, K.; Pridéaux, C.T.; Johnson, M.A.; Sheppard, M.; Fahey, K.J.
Virology 184, 404-410, 1991
A;Title: Nucleotide sequence of the gene encoding infectious laryngotracheitis virus gly
A;Reference number: A40567; MUID:91335774; PMID:1840710
A;Accession: A40567
A;Molecule type: DNA
A;Residues: 1-883 <KON>
A;Cross-references: GB:M64927
C;Superfamily: herpesvirus glycoprotein B
C;Keywords: glycoprotein; transmembrane protein
F;1-26/Domain: signal sequence #status predicted <SIG>
F;27-883/Product: glycoprotein B #status predicted <GPB>
F;725-744/Domain: transmembrane #status predicted <TM1>
F;751-771/Domain: transmembrane #status predicted <TM2>
(F;40,102,121,211,262,360,579,635,649/Binding site: carbohydrate (Asn) (covalent) #status
```

```
Query Match 18.4%; Score 836; DB 1; Length 883;
Best Local Similarity 27.9%; Pred. No. 9e-42;
Matches 222; Conservative 158; Mismatches 308; Indels 108; Gaps 23;
Qy 57 SHGERGNENRD--SEQNKNIY-----GSPSTFPYR-----VCSASGVGDVFRFQTDHV 104
Db 34 SHGTWNNSPHDTASMDVGKISFSEAIGSGAPKEPQIRNIFACSSPTGASVARLAQPRH 93
Qy 105 CPDASDMVH-SEGILLIYKQNIIPFMRVRKYRKYVWTSTVYNGIYSDSIINQHTFYKSI 163
Db 94 CHRHADSTNNTEGTAIVFKQNIADYVNVTLYKHITVTWALFSPQITNEVTRVPI 153
Qy 164 EPWETKMDTIYQCNLSRLNTGNGLLTYDRDDINMTVFLQP-----VDGVTPDV 214
Db 154 DYHEIVRIDRSGECSSKATYKHNMFPEAYDNDEAEKKLPLVPSSLRSTVSKAFHTTNT 213
Qy 215 KRYGQPELYLEPCWFGWSYRRRTVNCELMDMARSNPDPFVTATGTATVEMSPWSG 274
Db 214 KRHOTL-----GYRTSTSDCVVEYLQARSVPYDYFGMATGTVEISPPYT- 260
Qy 275 EDDHENKMKHPFVSVINNYK---VVDYQNR---GTVPGLGKTRIFLDRBEVTLSEKH 327
Db 261 -----KNTGPRHRSVYRDYRFLFIANYQVRDLETQIRPPKRNFLTDEQFTIGDAM 314
Qy 328 LKNMSYCLTLWKAIFYNGIOIEHSGSYHFVANDITASFTTSKEDKMFNTT-----YHCL- 382
Db 315 EEKESVCTLSKWIEVPEAVRVSYKNSYHFSCLKDMTMTFFSGKQ---PFNISRHLHAECPV 371
Qy 383 ---NEEIKAEIKKYAKVNSTHXYKGLKYPKTDGGLYLVWQPLIQNRLLD---AKNKLN 436
Db 372 TIASEAIDGIFARKY---SSTHVRSGDIEYVLGSGGFLIAFOKLMSHGLAEYLEEAQRQ 428
Qy 437 NETYSRRSROAESTTDPMMETGNGAGGESNSITVAQOVAYDNLRIRINNILEDL 496
Db 429 NHLPRGRERRQAAGRRTASIQ--SGPOGDRITTHSSATFAMLPAYDKIQAHWNELIGNL 486
Qy 497 SKAWCREQHRALVWNLKINPTSVMSIYNRPVSARIGDVISVSNCTIWWQTSVSLH 556
Db 487 LEAWCELQNRQLIIVHEMKLNPNLSMTSLFGQPV SARLLGDIIVAVSKCIEIPIENIRMQ 546
Qy 557 KSLRLLSASDEKCFSRPPTVKFM-----NDSTIYKQGLGVNNEILLTITTYL 603
Db 547 DSMR-VPGDPTMCTYRPLVIFRYSSSPESQFSANSTENHNLGILGQGEHNEILQGRNLI 605
Qy 604 ETQCENTEYFQAKTDMYIKVYEBHLKTVPLSSITTLDTFIALNFTLLENVDEKVIELYT 663
Db 606 EPCMINHRRFLLGENVILLVEDYTFVRQVASELEEVSTFINLNATILLEDLDFVPEVYI 665
Qy 664 RDEKRLSNVFDIETMFREYNIAQR-----VSLGRKDLDDLTNRNQFVDAFGLMDD 716
Db 666 REELRDTGTLNDDVRYQNIYKRNPRPDIDTVIRGDRGDAL-----PRAIADFFGNLTGE 720
```

```
Qy 717 LG-AVGQTVNAVSGVATLPSFSSIVTGFINFIPKPNFGCM-----LMIIVIGLVFAIYFLT 770
Db 721 VGLAUGTIVMTAAAVI-----STVSGIASFLSNPFAALAIAGIAVIVSVIILGLLAFKVMN 776
Qy 771 KTKIYETAPIKMIYP 786
Db 777 LKSN-----PVQVLEP 787
```

```
RESULT 26
VGBECS
glycoprotein B precursor - equine herpesvirus 1 (strain Ab4p)
C;Species: equine herpesvirus 1
A;Note: host Equus caballus (domestic horse)
C;Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 16-Jul-1999
C;Accession: G36798
R;Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.
submitted to GenBank, March 1992
A;Description: The DNA sequence of equine herpesvirus-1.
A;Reference number: A36805
A;Accession: G36798
A;Molecule type: DNA
A;Residues: 1-980 <TEL>
A;Cross-references: GB:M86664; NID:G330791; PIDN:AAB02468.1; PID:G330825
R;Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.
Virology 189, 304-316, 1992
A;Title: The DNA sequence of equine herpesvirus-1.
A;Reference number: A41831; MUID:92295566; PMID:1318606
A;Contents: annotation; possible protein-coding frames
A;Note: neither amino acid nor nucleotide sequence is given
C;Genetics:
A;Gene: 33
C;Superfamily: herpesvirus glycoprotein B
C;Keywords: glycoprotein; transmembrane protein
F;1-85/Domain: signal sequence #status predicted <SIG>
F;67-76/Domain: transmembrane #status predicted <TM1>
F;86-980/Product: glycoprotein B #status predicted <WAT>
F;850-868/Domain: transmembrane #status predicted <TM2>
F;165,275,380,423,497,514,515,560,727,749,952,971/Binding site: carbohydrate (Asn) (coval)
```

```
Query Match 18.4%; Score 835; DB 1; Length 980;
Best Local Similarity 26.9%; Pred. No. 1.2e-41;
Matches 250; Conservative 161; Mismatches 383; Indels 134; Gaps 31;
Qy 32 VTSPPNTATWSTESPLTGHYTHDSHGSRGNENRDSSEQNKNIYVGSSTFPYRVC-SA 90
Db 83 VRAVPTTSPPTSTPTS--MSTH--SHG-----TVDPTLLPTETPDLRLAVRE 127
Qy 91 SGV-----GP-----VRFQTDHVCPDAS-DMVHSEGLLIYKQNIIPFMRVRK 134
Db 128 SGLAEDGDFYTCPPPTGTSTVVRIBPPTCPKPDGRNFTGIAVIFKNIADYKFRANV 187
Qy 135 YRKVVTTSTVANGIYSDSIINQHTFYKSIETPETER-----MDTIYOCFNS---LRLNTGG 187
Db 188 YKDIIVTVRWKGYSHTSLSDR---YNDRVPSVSEIFGLIDSKGCKSSKAEVLR----D 240
Qy 188 NLITYV---DRDINMTVFLQPDVGTVPDKRYGSOPELYLEPGWF-WGYSRRRTTNCE 243
Db 241 NIMHAYHDEDEVELD--LVPSKFPATPGARAWQTTNDTTSYVGMMPWRHY-TSTSVNCI 297
Qy 244 LMDMPARSNPDPFFVTATGDTVENSPPFW-----SGEDDENKMKHPFVSVINNYKVD 299
Db 298 VEEVARSVYVDSFALSTGDIIVASFPFYGLRAAARIENHNSVAQER---FRQVEGRPRD 354
Qy 300 YQNRGTVPLGKTRIFLDRBEYTLSEKHLKNNMSYCPCLTWKAFYNGIQIEHSGSYHFVAN 359
Db 355 LUSKLAQEPVTKNFITTHVTVSNWNTKKVEACTLTWKVEVDLVRDEFRCGSYFTIR 414
Qy 360 DITASPTTSKEDMK-EFNTTYHCLNEEIKAEIKKYAK-VNSTHXYKGLKYPKTDGGLY 417
Db 415 SISSTPISNTQPKLESAPLTCVSKAEKAIDSIIYKQYESTHVPFSGDVEYVYIARGGFL 474
```

```
QY 418 LVMQPLI-----QNRLLDAKNKL-----NNETYGRRRRRQAESTTDDMMMENT 459
Db 475 IAFRPMLSNELARLYLNLVRSNRTYDLKNLLNPNANNNTTTRRRSLLSVPEQPQTD 534
QY 460 G-----NGAGGEYSSE-----NSITVAQVQAYADNLRIRINNI 492
Db 535 GVHREOILHRLHRAVEATAGTDSNVVTAQLELTKTSSIBFAMLOFAYDHIQSHVNM 594
QY 493 LEDLSKAMCREQRAALVNNELSKINPTSVMSMINRNPVSARKIGDVISVSNICVVDDOTS 552
Db 595 LSRIATAWCTLQNKERTLWNEWKINPSAIVSATLDERVAARVLDGVDAITHCAKI-EGN 653
QY 553 VSLHSLRLLSASDEKCFSPVPVTF---KPMNDSTIYKQGLGVNNEILLTTTTYLETQEN 609
Db 654 VYLQNSMR--SMDNSTCYSRPVPVTFITTKNANRSGIEGOLGEENEIFTERKLIPECALN 711
QY 610 TEYFQAKTDMYIKNYEHLKVPGLSSITTLDTFALNFTLLENVDFKVIELYTRDEKRL 669
Db 712 QKRYFKQKVEYVYENYTFVRKVPPEIEVISTYVELNLTLLEDRFLPLEVYTRAELD 771
QY 670 SNVFDIETMFREYNYAORVSGRLKDLDDLSTNRNQFVDAFGSLMDDLGAVGOTVNVAVS 729
Db 772 TGLLDYSEIQRNQHLARFYDI--DSVVNDNTAVIMQGIASFPGKLGKGEAVGTIVL 829
QY 730 GVATLFSSIVTGFINFKNPPCGMLMIIVIGVLFAYFLTKTKIYETAPIKMIYPRID 789
Db 830 GAAGAVSVTVSGIASFPLNPFGLAIGLLVIAGLVAFAFYRVVQIRSNPMKALYPITT 889
QY 790 KLKEREGKSEIAPIS-----BELERIVLAMIHQONSHMETKTRKDPKDS--- 835
Db 890 KALKNKATSYQONBEDGSDFDEAKLEAREMIKYMVSALERQEKKAIK--KNSGVG 947
QY 836 -ILTRAQNMLKRSG--YSNLKNAESVE 860
Db 948 LIASNVSKLALRRGPKVTRLOQNTME 975

RESULT 27
VGBESA
Glycoprotein B precursor - simian herpesvirus SA8 (strain B264)
C:Species: simian herpesvirus SA8
C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 16-Jul-1999
C:Accession: JQ1332
R:Bochers, K.; Weigelt, W.; Buhk, H.J.; Ludwig, H.; Mankertz, J.
J. Gen. Virol. 72, 2299-2304, 1991
A:Title: Conserved domains of glycoprotein B (gB) of the monkey virus, simian agent 8, i
A:Reference number: JQ1332; MUID:91374035; PMID:1895066
A:Accession: JQ1332
A:Molecule type: DNA
A:Residues: 1-885 <BOR>
A:Cross-references: EMBL:X56935; NID:G60438; PIDN:CAA40256.1; PID:G60439
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein; transmembrane protein
F:1-28/Domain: signal sequence #status predicted <SIG>
F:129-885/Product: glycoprotein B #status predicted <MAT>
F:712-732/Domain: transmembrane #status predicted <TM1>
F:737-752/Domain: transmembrane #status predicted <TM2>
F:760-780/Domain: transmembrane #status predicted <TM3>
F:68,122,379,411,659/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:197-558,114-514,188-252,345-393,581-618/Disulfide bonds: #status predicted

Query Match
Best Local Similarity 27.1%; Pred. No. 1.4e-41;
Matches 253; Conservative 164; Mismatches 392; Indels 124; Gaps 26;

QY 5 LKLRG---SVIALWLYQVALYSLSIAETGVTPSPNTATWSTESPLTGHYTHDSHGER 61
Db 1 MRPRGTPTSPFLPLVLLALV---IAAGRAAPAAAAAPTADPAATPALPEDEBVPD 56
QY 62 GNNENRDEEQNKNIYGSPTF-----PVRVCSASGVGVFRFQDTHVCPD 107
Db 57 GEGVATPAPAANASVEAGRATLREIKARDGDATFYVCPPTGTATVQVQEPQPCPR 116
```

```
QY 108 ASD-MVHSEGILLIYKQNIIPMFVRVRKYRKVVTSTV-----YNGIYSDSITNOHT 158
Db 117 APDQNYTEGIAVVFKENIAPYKFKATMYKDVTVSQVWFGRYSQFNGIED----- 169
QY 159 FYSKIEPWETEKMTIYO---CFNSLRLLNTGNNLLTYVDRDINNMTVFLQPDVGVTPDVK 215
Db 170 --RAPVPPE-EVMOKINAKGVCRSTAKVVRNNMESTAFHRDDHESDMALKAKAAT-RTS 225
QY 216 RYGSQPELYLPBGFWSYRRRTTVNCELMDMFARSNPPDFPVFTATGDTVEMSPFWSGE 275
Db 226 RGWHTTDLUKNPAPVEAPHRYGTTVNCIIVEEARSVPYDFEVLATGDFVMSPFYGYR 285
QY 276 D-DH-ENKMKHEKPFVSVINNY-KVVDYQNRGTVPGLKTRIFLDREYTLWSKEHLKMS 332
Db 286 DGSHEHTAYAADRFQVDGYVERDLSGRRAAAPV--TRNLLTTPKFTVGMWDWAPKPS 343
QY 333 YCPLTLWKAFYNGIQTEHSGSYHFVANDITASFTT-----SKEDMKFNTVYHCLNEE 385
Db 344 VCTLTKREVDEMRAEYGPSPFSSAALSTTFTANRTEYALSRLVDLAD-----CVGRE 397
QY 386 IKAETEKY-AKVNSTHSKYGLDYKFKTDGGLYLVMQPLIONRLLDKNKLNNEYSRSS 444
Db 398 AREAVDRIFLRVNGTHVKVQVYYLATGFLIAYQPLLSNALV-----ELYVREL 449
QY 445 RROAESTTDPMMMTGNGAGGE-----YSSSENSITVAQVQAYADNLR 486
Db 450 VR-----EQTRRPAGDGPCEAATPGPSVDPSPVERIKTSSVEFARLQFTYDHIQ 499
QY 487 IRINNILEDLSKAWCRQHRALVNNELSKINPTSVMSMINRNPVSARKIGDVISVNCI 546
Db 500 RHVNDMLGRIATAWCELQNRELTLWNEARRLNPGAISATVGRRRVSARMGLDVMVAVSTCV 559
QY 547 VVDQTSVSLHKSRLLSASDEKCFSPVPVTFKPMNDSTIYKQGLGVNNEILLTTTTYLET 606
Db 560 PVAPDNVIMQNSIG-VAARPGTCYSRPLVSPRYEADGGLVEGEGEDNEILERDALPCC 618
QY 607 QENTFYFQAKTDMYIKNYEHLKTVPLSSITTLDTFIALNFTLLENVDFKVIELYTRDE 666
Db 619 TVGHRRYFTFCAGYVYFEEYAYSHQGLRADVTVTVSTFINLNLTMLEDHEFPVLEVYTRQE 678
QY 667 KRLSNVFDIETMFREYNYAORVSGRLKDLDDLSTNRNQFVDAFGSLMDDLGAVGOTVYN 726
Db 679 IKDSGLLDYTEVQRRNQHLARFADI-DTVIKADAAHALFAGLY-SFPEGGLDVGRAVGK 736
QY 727 AVSGVATLFSIVTGFINFKNPPCGMLMIIVIGVLFAYFLTKTKIYETAPIKMIYP 786
Db 737 VVMGIVGVSVAVSGVSVSFLSNPFGALAVGLLVLAGLAAAFAFPFRYVNRQLQNPMLALYP 796
QY 787 EIDKLKEREGKSEIAPIS-----EEELERIVLA-----MHIHQONSHMETKTRKDP 832
Db 797 ---LTTKELKSDCAPLAGGEDGAEFDDEAKLAQAREMIRYMALVSAMERTEHKARKG 852
QY 833 KDSIL----TRAQNMLKRSGYSNLKNAESVEM 861
Db 853 TSALLSAKTVADVMRKARPRYSPLRDTDBEEL 885

RESULT 28
VGBE31
Glycoprotein B - human herpesvirus 3
N:Alternate names: Glycoprotein II
C:Species: human herpesvirus 3, varicella-zoster virus
C:Date: 30-Sep-1998 #sequence_revision 30-Sep-1998 #text_change 16-Jul-1999
C:Accession: E27214
R:Davidson, A.J.; Scott, J.E.
J. Gen. Virol. 67, 1759-1816, 1986
A:Title: The complete DNA sequence of varicella-zoster virus.
A:Reference number: A27345; MUID:86306657; PMID:3018124
A:Accession: E27214
A:Molecule type: DNA
A:Residues: 1-868 <DAV>
A:Cross-references: EMBL:X04370; NID:G59989; PIDN:CAA27914.1; PID:G60020
C:Genetics:
```

A;Gene: 31

C;Superfamily: herpesvirus glycoprotein B

C;Keywords: glycoprotein; transmembrane protein

F;697-713/Domain: transmembrane #status predicted <TM1>

F;726-742/Domain: transmembrane #status predicted <TM2>

F;84,194,372,416,440,494,557,623,781,836,864/Binding site: carbohydrate (Asn) (covalent)

Query Match 18.3%; Score 832; DB 1; Length 868;

Best Local Similarity 26.7%; Pred. No. 1.5e-41;

Matches 236; Conservative 168; Mismatches 365; Indels 116; Gaps 27;

Qy 40 TWSTESPLTHGYGTHDSHGCGNNENRDSSEQNKNIYGPSTPPYVCSAGSGVDVFRF 99

Db 23 TQSEDITSAHLGQD-----EIREAHKQDAETKPT--FYVCPPTGCTIVRL 70

Qy 100 QTDHVCDDPDMVH-----SEGILLIYKQNIIPMFVRKYRKVVTTSTVNGIYSDSIT 154

Db 71 EPTRTCPD-----YHLGNFTGEGIAVVKENIAAYKFKATVYKDVIVSTAWAGSSYQTIT 126

Qy 155 NOHTFYKSIPEWE--TEKMDTIYQCFNSLRNTGCGNLLTYV-----DRDDINMTV 202

Db 127 NRYADRPVPIVSEITDIDKFGKC-----SSKATYVRNNHKVAFNEKNQPMPL 177

Qy 203 FLOPVDGVTDPVKRYGSOPELYLEPGFMWGSYRRRTTVCNCELMDFARSNPPDFVAT 262

Db 178 IASKVNSV--GSKAWHTTNDIYMWAG--TPGYRTGTSVNCIIIEVEARSIFPYDSFGLST 234

Qy 263 GDTVEMSPFWGSD-----DHEN-----KMHKPMFVSVNNYKVVYDYNQRTGTVPLGKTRIF 314

Db 235 GDIIYMSPPFGLRDGAYREHSNAYAMDRFHQ-----FEGYRQRDLDTRALLE--PAARNF 286

Qy 315 LDREYTLSEKHLKNNKSYCLTLWKAFYNGICQTEHSGSYHFVANDITASFTTSKEDMKE 374

Db 287 LVTHPLIVGNWKKPRTEVCSLVKRWEDVVRDEYAHNFRFTMKTLSFTFIS----ETNE 343

Qy 375 FNTTY-----HCLNEIEIKAEIEKKY--AKVNSTHSGYGLKFKYTDGGLYLVWQPLIQNRL 429

Db 344 FNLQIHLSCQCKEERAILNRIYTRYNSHVTGDIQIYLARGGVVFPQLLSLSLA 403

Qy 430 -----DAKNLNNETYRRRRQRAESTTDPMMETNGAGGEYSSENSITVAQVQYAYDN 484

Db 404 RLXLQELVRENTNHSPOKHPTNTRSRSPVVELRANRT---ITTTSSVEFAMLQFTYDH 460

Qy 485 LRIRINNILEDLSKAWCREQRAALVWNLKINPTSVMSMIYNNRPVSAKRIGDVISVN 544

Db 461 IOEHVNEMLARISSWCOLQNRERALSGLFPINPSALASTILDQRYKARILGDVVISVN 520

Qy 545 CIVV--DOTSVSLKSLRLLSASDEKCFSPRPVTFKFMNDSTIYKQGLGVANNEILLTTYL 603

Db 521 CPGLGSDTRIILQNSMR--VSGSTTRCYSRPLISIVSLNGSGTVEGQLGTDNELIMSRDLL 579

Qy 604 ETCQENTYFFQAKTDMYIYKNYEHKLTVPSSITTLDTFIALNFTLLENVDFKVIELYT 663

Db 580 EPCVANHKRYPLFGHHVYVYEDYVREIAVHDVGMISTVVDLNLTLKKDREFPMLQVYT 639

Qy 664 RDEKRLSNVPDIEFMFNEYNYAORVSGRLKDL--LDLSTNRNQFVDAFGSLMDLGVGQ 722

Db 640 RDELRTGLLDYSEIQRRNQMHSLURFYDIDKVQVYDSGT---AIMQMAQVFOGGLGTAGQ 696

Qy 723 TVNNAVSGVATLFSIVTGFNFINKPFGGMLMIIVIGVLFALYFLTKTKIYETAPIK 782

Db 697 AVGHVVLGATCALLSTVHGFTTFLSNPFGALVGLLAVGLVAFAFYAYRVLKLTSPMK 756

Qy 783 MIYPEIDK--LKE--REGKSEIA-----PISE-----EEERIV 812

Db 757 ALXPPLATTKGLQLPEGMDFPAEKNATDTPIEIGDSQNTSPSVNSGDFDPKFEAQEMI 816

Qy 813 LAMHIHQONSHMETKTKRDKPSLITRAQ---NWLKRSCYSNLK 854

Db 817 KYMTLVSAAERQESKARKNKTSALLTSRLTGLALNRNRGYSRVR 861

RESULT 29 ,

VGBEIL

glycoprotein B precursor - infectious laryngotracheitis virus (strain Thorne V882)

C;Species: infectious laryngotracheitis virus

A;Date: host Gallus gallus (chicken)

C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 16-Jun-2000

C;Accession: A38478; JQ0951

R;Griffin, A.M.

J. Gen. Virol. 72, 393-398, 1991

A;Title: the nucleotide sequence of the glycoprotein gB gene of infectious laryngotracheitis virus

A;Reference number: A38478; MUID:91132136; PMID:1847176

A;Accession: A38478

A;Molecule type: DNA

A;Residues: 1-883 <CRI>

A;Cross-references: GB:D00818; NID:G221905; PIDN:BAA00699.1; PID:G221907

C;Comment: This transmembrane protein is essential for viral infectivity.

C;Genetics:

A;Gene: gB

C;Superfamily: herpesvirus glycoprotein B

C;Keywords: glycoprotein; transmembrane protein

F;1-26/Domain: signal sequence #status predicted <SIG>

F;27-883/Product: glycoprotein #status predicted <GCP>

F;725-744/Domain: transmembrane #status predicted <TM1>

F;751-771/Domain: transmembrane #status predicted <TM2>

F;102,121,211,262,360,579,635,649/Binding site: carbohydrate (Asn) (covalent)

Query Match 18.3%; Score 830; DB 1; Length 883;

Best Local Similarity 27.8%; Pred. No. 2e-41;

Matches 221; Conservative 159; Mismatches 308; Indels 108; Gaps 23;

Qy 57 SHCERGNENRD--SSEQKNIY-----GSPSTFFVR---VCSAGVGVDVFRFQTDHV 104

Db 34 SHGIAGIIPDRDTASMDVKISFSEALGSGAPKEQIRNIRIFACSSPTGASVARLAQPRH 93

Qy 105 CPDASDMVH--SEGILLIYKQNIIPMFVRKYRKVVTTSTVNGIYSDSITNQHFTFYKSI 163

Db 94 CHRHADSTNNMTEGIAVVFQNIAPYVFNVTLYYKHITTTTMMALFSRQITNEVYVTRVPI 153

Qy 164 EPWETERMDTIYQCFNSLRNTGCGNLLTYVDRDDINMTVFLQP-----VDGVTDPDV 214

Db 154 DYHEIVRIDRSGECSSKATYHKNFMFPEAYDNDEAEKLPVPSLLRSTVSKAFHTTNTFT 213

Qy 215 KRYGSOPELYLEPGFMWGSYRRRTTVCNCELMDFARSNPPDFVATGDTVMSFWSG 274

Db 214 KRQTL-----GYRTSTVDCVVEYLQARSVPYDFGMATGDTVEISPFYT- 260

Qy 275 EDDHENKMHKPMFVSVNNYK---VVDYQNR---GTVPLGKTRIFLDREYETLSWEKH 327

Db 261 -----KNTTGPRRHSVYRDYRFLEIANYQVRDLETGQIRPPKRNFLTDEQFTIGWDAM 314

Qy 328 LKNMSYCPITLWKAFYNGICQTEHSGSYHFVANDITASFTTSKEDMKEFNTT-----YHCL- 382

Db 315 BEKESYCTLSKRIEVEPAVRVSYKNSYHFSKLDMTMTTFSSGKQ---PFNISRHLAECPV 371

Qy 383 ---NEEIKAEIEKKYAKVNSTHSGYGLKLVFKYTDGGLYLVWQPLIQNRLD---AKNKLN 436

Db 372 TIASEALDGIIPARKY---SSTHVRSGDIEYILGSGGFLIAFOKLMHSHGLAEYLEBAQRO 428

Qy 437 NETYSRRSRQRAESTTDPMMETNGAGGEYSSENSITVAQVQYAYDNLRIRINNILEDL 496

Db 429 NHLPRCERRRQAGRRRTASLQ--SGPQGDRIITHTSSATFAMQLQAFYDKIQAHVNELIGNL 486

Qy 497 SKAWCREQRAALVWNLKINPTSVMSMIYNNRPVSAKRIGDVISVNCIIVDQTSVSLH 556

Db 487 LEAWCELRQRLQIVHMEKMLNPSLMTSLFGQPVARSALLGDIVAVSKCIPIENIRMQ 546

Qy 557 KSLRLLSASDEKCFSPRPVTFKFM-----NDSTIYKQGLGVANNEILLTTTTL 603

Db 547 DSMR--WPGDPTWCYTRPVLIFRYSPPESQPSANSTENHNDILGQGLGHEHNEILQGRNLI 605

Qy 604 ETCQENTYFFQAKTDMYIYKNYEHKLTVPSSITTLDTFIALNFTLLENVDFKVIELYT 663

Db 606 EPCMINHRRYFLLGENYLLYEDYTFVRQVNASEIBEVSIFINLNTAILEDLDLPVPVEVYT 665

Qy 664 RDEKRLSNVFDIETMFREYNYAQR-----VSLRKKOLLDLSTNRNQFVDAFGSLMDD 716
Db 666 REELRDGTLLNYDDVVRQYQYINRYKRRIDITVIRGDRDAI-----FRAIADFFGNTLGE 720
Qy 717 LG-AGVGTQVNVAVSGVATLFSSTVCTFINIKNPGGM-----LMIIVIGVLFAIYPLT 770
Db 721 VGKALGTQVNVAAAVI-----STVSGIASFLSNPFAALGIGIAVAVVSIILGLAFKYVMN 776
Qy 771 KTKKIYETAPIKMIYP 786
Db 777 LKSN-----PVQLFEP 787

RESULT 30
S26690
glycoprotein B - infectious laryngotracheitis virus
C:Species: infectious laryngotracheitis virus
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 07-May-1999
C:Accession: S26690
R:Poulsen, D.J.; Adams Burton, C.R.; O'Brian, J.J.; Rabin, S.J.; Keeler Jr., C.L.
A:Title: Identification of the infectious laryngotracheitis virus glycoprotein gb gene
A:Reference number: S26690; MUID:92180317; PMID:1665614
A:Accession: S26690
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-873 <POU>
A:Cross-references: EMBL:X56093
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein; transmembrane protein

Query Match 18.2%; Score 827; DB 2; Length 873;
Best Local Similarity 27.8%; Pred. No. 3e-41;
Matches 221; Conservative 158; Mismatches 309; Indels 108; Gaps 23;

Qy 57 SHGERGNENRD--SEQNKNIY-----GSPSTFPYR-----VCSASGVGDVFRFQDHV 104
Db 24 SHGIAGIDPRDTASMDVGKISFSEAIGSGAPKEPQIRNRIFACSSPTGASVARLAQPRH 83
Qy 105 CPDASDMVH--SEGILLIYKQNIIPFMRVKYRKVVTSTVYNGIVSDSITNOHTFYKSI 163
Db 84 CHRHADSTNWTEGIAVVFQNTIAPVNTLYKYHITVTTWALFSRPOITNEYVTRVPI 143
Qy 164 EPMETEKMDTIYOCFNSRLNTGNNLLTVDRDDINNTVFLOP-----VDGVTPDV 214
Db 144 DYHEIVRIDRSGECSSKATYKHNMFPEAYDNDEREKKLPLVPSLLRSTVSKAFHTNFT 203
Qy 215 KRYGSOPELYLEPGNFWGWSVRRTTVNCELMFARSNPPFFVTVATGDTVEMGPFWSG 274
Db 204 KRHQTL-----GYRTSTSDCVVEYLQARSVYPDYFGMATGDTVEISPFVT- 250
Qy 275 EDDHENKMHKEKPMFVSINNYK---VDYQNR---GTVPGLKTRIFLDREBYTSLWEKH 327
Db 251 -----KNTGPRRHSVYRDYRELEANTVOVDLETGQIRPKKRNFLTDEQFTIGWDM 304
Qy 328 LKMSYCPLTLKAFYNGIQTESGSHYFVANDITASFTTSKEDMKEFNTT----YHCL- 382
Db 305 EEKESVCTLSKMEVPEAVRVS YKNSYHFSKDMTWTFSSGKQ---PFNISRLHAEVCP 361
Qy 383 ---NEEKABIEKKYAKNSTHSGYGLKYFKYTDGGLYVWQPLQNRLLD---ANKLN 436
Db 362 TIATEAIDGIFARKY---SSTHVRSGDIEYLGSGGFLIAFKLMSHGLAEMYLEAQRQ 418
Qy 437 NETYSRRSRQARESTTDPNMWMTGNGAGGEYSSENSITVAQVQYAYDNLRIINNILEDL 496
Db 419 NHPGRGRERQAAGRRTASIQ---SGPOGDRITHTSSATFAMLOQFAYDKTQAHVNELIGNL 476
Qy 497 SKAWCREQRAALVWNELSKINPTSVMNMYNRPVSAKRIGDVISVNCIVVDQTSVSLH 556
Db 477 LEAWCELONQLIWHMKLNPNLSLTFPGQVSARLLGDIVAVSKCIEPIENIRMQ 536
Qy 557 KSLRLLSASDEKCFSRPPVTFKPM-----NDSTIYKQGLGVNNEILTTTLYL 603
Db 557 KSLRLLSASDEKCFSRPPVTFKPM-----NDSTIYKQGLGVNNEILTTTLYL 603


```
Db 467 YLARGGELIAPRMLSNELARLYNELVRSNRTYDLKNLLNPNANNNNNTTTRRRSLVS 526
Qy 452 TDPMMWTG-----NGAGGESSE-----NSITVAQVQYAYDN 484
Db 527 PEPOPTQGVHREQIHLRLHKRAVEATAGTDSNNVTAKOLELIKTTSSIEFAMLQFAYDH 586
Qy 485 LRIRINNILEDLSKAWCREQRAALVMNELSKINPTSMVSMIYNRPYSAKRIGDIVSVSN 544
Db 587 IQSHVNEMLSRIATAWCTLQNKERTLWNEVKINPISAIVATLDERVAARVLGDVIAITH 646
Qy 545 CIVVDQTSVLSHKSRLLSASDEKCFRPPVTF---KPMNDSTIYKQGLVNNIEILLTTT 601
Db 647 CAKI-EGNVYLQNSMR--SMDSNCTCYSRPPVTFITTKNANNRGSIEQGLGEENEIPIERK 703
Qy 602 YLETCQENTYFFOAKTMYIKNYEHLKTVPLSSITLDTFIALNFTLLENVDFKVEL 661
Db 704 LIIEPCALNQKRYFKGKEYVYENYTFVRKVPPTIEIVISTYVELNLTLEDREFFLEV 763
Qy 662 YTRDEKRLSNVFDIETMFREYNYAQRVSLGRKDLLDLSNTRNQFVDAFGSLMDLGAVG 721
Db 764 YTRALEDTGLLDSEIQRRNQLHALRFPDI--DSVVNDNTAVIMOGIASFFKGLKVG 821
Qy 722 QTVVNAVSGVATLFSSTVITGFINKPFGGMLMIIVVIGVFAIYFLTKTKIYETAPI 781
Db 822 EAVGTLVLGAAGAVVSTVSGIASFLNPNFPGGLAIGLLVIAGLVAAPFAYRYVMQIRSNPM 881
Qy 782 KMIYPEIDKLKERECKSEIAPISE-----ELEIRIVLAMHHQONSHMETYRK 830
Db 882 KALYPIITTKANKAKTSYQONEEDGSDFDEAKLEBAREMIKYMVMVALEXQEKKAIK 941
Qy 831 DPKDS----ILTRAQNMLRKSG--YSNKLNAESVE 860
Db 942 --KNSGVGLIASVSKALARRGPKYTELQONDME 975
```

RESULT 32

```
B48349
Glycoprotein B precursor - ateline herpesvirus 1 (strain Lennette)
C;Species: ateline herpesvirus 1
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 09-Sep-1994
C;Accession: B48349
R;Eberle, R.; Black, D.
Arch. Virol. 129, 167-182, 1993
A;Title: Sequence analysis of herpes simplex virus gB gene homologs of two platyrrhine
A;Reference number: A48349; MUID:93228440; PMID:8385913
A;Accession: B48349
A;Molecule type: DNA
A;Residues: 1-933 <EBE>
A;Note: sequence extracted from NCBI backbone (NCBIN:129063, NCBI:P:129065)
C;Superfamily: herpesvirus glycoprotein B
C;Keywords: Glycoprotein; transmembrane protein
F;1-35/Domain: signal sequence #status predicted <SIG>
F;36-933/Product: glycoprotein B #status predicted <MAT>
F;775-794/Domain: transmembrane #status predicted <TMI>
F;801-818/Domain: transmembrane #status predicted <TM>
F;107,161,418,450,697,747/Binding site: carbohydrate (Asn) (covalent) #status predicted
```

```
Query Match 18.1%; Score 822; DB 1; Length 933;
Best Local Similarity 26.4%; Pred. No. 6.6e-41;
Matches 224; Conservative 164; Mismatches 378; Indels. 84; Gaps 20;
```

```
Qy 30 TGVTSPTATWSPESPLTGHYCHTHDSSHGKGNV---ENRDSSEQNKNIYSPSTFPY 85
Db 80 SGPESP-----GDRPRPGGGSGRRRSGPNGTSAARRQLRESLRIOAEYAAAF 133
Qy 86 RVCSASGVGVFRFQTDHVCVD-ASDMVHSEGLLIYKQNIIPFMFVRKYRKVVTSTV 144
Db 134 YVCPPTGATVQVEERPCDPAAGKNFTGIAVVFENIAPVKFTATKYKEITVSQT 193
Qy 145 Y-----NGIVSDSTNQHTFYKSIETPWE--TEKMDTIYQCFNSURLNTGNLLTYVD 194
Db 194 WQGSRYLQLTGLYND-----RAPVPFSEITDLINGKRCRSDVYTRSORRTAYD 244
```

```
Qy 195 RDDINMTVFLQVGDVTPDKRYGSOPELYLEPGWFWSYRRRTTNCCLMDMFARSNPP 254
Db 245 GDEWGEVALVPAKTSPTNSRGWYTTDRVY-ANAHAGFYKTGTTVNCIVVEEMEARAFP 303
Qy 255 FDFVFATGDTVMSFPWSGEDD--HENKMHKXPWFVSVINNY-KVVDYQNRGTVPGLKT 311
Db 304 YDSEVLATGEFVYASFPSGFSDARERNRYAPDRQROVDFYPRDLDSQRAATPV--V 361.
Qy 312 RIFLDREBYTLSEKHLKNMSYCPFLWKAPYNGIQTEHSGSYHFVANDITASFTSKED 371
Db 362 RNULTTPTFTGVGDMKPKRPNVCSVTWKQVVEEMVRAEYGSAPRFTSAALSATFTS 417
Qy 372 MKEFNITYH-----CLNEEIKAEIEKKYA-KVNSTHSKYGDLKYFKTDCGLYLWQ 421
Db 418 ----NUTQPPPELIEHSDCVAREAAESI EAIYARRYNASHVRVGGVYLLAAGFLLAQ 473
Qy 422 PLIQNRLLDAKNKLNNETYSRRSRQAEESTTDPMMEMTNGAG--GEYSSENSITVAQVQ 479
Db 474 PLLSNSLAEMYR---EALLGRSGDLAALAPPVAPASGAGPRGTISTTQVFEARLQ 530
Qy 480 YADNLRIRINNILEDLSKAWCREQRAALVMNELSKINPTSMVSMIYNRPVSAKRIGDV 539
Db 531 FTYDHIQKHVNEMLGRIAAAWCOLQOELVLMNEARKLNPAGIASATVGTVRGARM LGDV 590
Qy 540 ISVSNICIVVDQTSVLSHKSRLLSASDEKCFRPPVTEKFMNDSTIYKQGLVNNIEILLT 599
Db 591 MAVSTCIPSPDNVIMQNSMR-IPGDPKTCYARPLVSFRYTDGELVEGQGLGDNELRLE 649
Qy 600 TTYLETQENTYFFOAKTMYIKNYEHLKTVPLSSITLDTFIALNFTLLENVDFKVI 659
Db 650 QNNVEPCTVGHKRYFVGGGFVYEEYAYSHQVSRAQVDPVWVSTFDLNLTMLEDHEFLPL 709
Qy 660 ELYTRDEKRLSNVFDIETMFREYNYAQRVSLGRKDLLDLSNTRNQFVDAFGSLMDLGA 719
Db 710 EYVTRREIKDGLLDYAEVQRNQLHALRFSIDIRIMND-SANA-ALMAGLARFFDGMGD 767
Qy 720 VGVTVNAVSGVATLFSSTVITGFINKPFGGMLMIIVVIGVFAIYFLTKTKIYETA 779
Db 768 AGAIGRAVLGVTEGLISVVSGVSFLSNPFGALAVGLLVLAAGLVAFAFMRIIMLRAN 827
Qy 780 PKIMYIPEI-----DKLKERECKSEIAPISEEELE-----RIVLAMHI 817
Db 828 PMRALYPLITTSIGIKAEARAALGSGGDKGAGDGAAGVEDFDEAKLEAARDMIRYMTLVA 887
Qy 818 HQONSHMETK 827
Db 888 MERTAHKAKK 897
```

RESULT 33

```
VGBEQH
Glycoprotein B precursor - equine herpesvirus 4 (strain 1942)
C;Species: equine herpesvirus 4
C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 16-Jul-1999
C;Accession: A31880
R;Riggio, M.P.; Cullinane, A.A.; Onions, D.E.
J. Virol. 63, 1123-1133, 1989
A;Title: Identification and nucleotide sequence of the glycoprotein gB gene of equine her
A;Reference number: A31880; MUID:89125704; PMID:2915378
A;Accession: A31880
A;Molecule type: DNA
A;Residues: 1-919 <RIG>
A;Cross-references: GB:M26171; NID:G341446; PIDN:AAA46106.1; PID:G514920
C;Superfamily: herpesvirus glycoprotein B
C;Keywords: glycoprotein; transmembrane protein
F;1-28/Domain: signal sequence #status predicted <SIG>
F;29-919/Product: glycoprotein B #status predicted <GPB>
F;740-809/Domain: transmembrane #status predicted <TM>
F;106,216,321,364,438,456,493,499,666,688/Binding site: carbohydrate (Asn) (covalent) #st
```

```
Query Match 18.0%; Score 819.5; DB 1; Length 919;
Best Local Similarity 26.5%; Pred. No. 9.1e-41;
```


Db 796 DNTAVIMQGIATPFGKLGVEAVGTVLGAAGAVVSTVSGIASFINNPGFLAIGLLVI 855
Qy 761 GVLFAIYFLPKTKIYETAPIKMIYPEI-----DKLERGKSEIAPISE-----BELE 809
Db 856 AGLVAFAFYVYVQLRSNPMKALYPITTRSLKNKAKASGYQNDDDTSDFDAKUEEAR 915
Qy 810 RIVLAMHIHQNSHMETKTRDKP--DSILTRAQNMUKRSG--YSNLKNAESVE 860
Db 916 EMIKYMSVSALEKQEKANKKNGVGLIASNVSKALRRGPKYTRLREDDPWE 970

RESULT 35
VGBBBH
glycoprotein B precursor - bovine herpesvirus 2 (strain BMV)
C;Species: bovine herpesvirus 2
C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 16-Jul-1999
C;Accession: C29242
R;Hammerschmidt, W.; Conraths, F.; Mankertz, J.; Pauli, G.; Ludwig, H.; Buhk, H.J.
Virology 165, 388-405, 1988
A;Title: Conservation of a gene cluster including glycoprotein B in bovine herpesvirus 2
A;Reference number: A94381; MUID:88306231; PMID:2841793
A;Accession: C29242
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-917 <HAM>
A;Cross-references: GB:M21628; NID:g330752; PID:AAA46053.1; PID:g330753
C;Superfamily: herpesvirus glycoprotein B
C;Keywords: glycoprotein; transmembrane protein
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-917/Product: glycoprotein B #status predicted <GPB>
F;578-594/Domain: transmembrane #status predicted <TM1>
F;770-786/Domain: transmembrane #status predicted <TM2>
F;795-811/Domain: transmembrane #status predicted <TM3>
F;48,110,164,278,421,453,505,564,692/Binding site: carbohydrate (Asn) (covalent) #status

Query Match 17.7%; Score 805; DB 1; Length 917;
Best Local Similarity 27.5%; Pred. No. 6.6e-40;
Matches 229; Conservative 164; Mismatches 335; Indels 106; Gaps 27;

Qy 74 KNYGSPSTPYRVCSAGVGDFRFTQDHVCPD-ASDMVHSGILLIYKQNIIPMFVRV 132
Db 125 RGLGSGDDPNFYVCPPTGATVVRLEPRPCPLKGLNFTGEGIAVTFKENLAPYKFA 184
Qy 133 RYKRVKVTTSVNGIYSDITNQHTFYKSIE-----PWETEXMDTIY---QCENSRLN 184
Db 185 TWYKAVTASVWSG-VS-----YNQFMNIFEDRAPIPFE-EIVDRIGHRGMCCLSTAK-- 235
Qy 185 TGGNLLTYVDRDDINMTVFLQPDGVTPDVVKRYGQPELYLEPCWF-----WG 232
Db 236 -----YV-RNNLETTAFNDAD--BHEMKLVPAESAPGLHGWHTTRLKNPTGSAW- 284
Qy 233 SYRRRTTVNCELMDMFARSNPPDFVTATGDTVEMSPFWSGEDDHENKMKHEKPFVSVI 292
Db 285 IHRHGTTVDCIVDEAKSGYPYNEFVLATGDFVYAPFF-GYRDSGSHSN-----AYAA 339
Qy 293 NNKVVQYQNGTVP-----LGK-----TRIFLDREBYTLSEKHLKNMSYCPILTWKAF 342
Db 340 DRFKQVD-----GFFPRDFGTRRHGSPVTYNLLTTPMTVGMWAPKRPVCTMTKREV 395
Qy 343 YNGIQTEHSGSYHFVANDITASFTT-----SKEDMKFNTTYHCLNEIEIKAEIKKY- 394
Db 396 PEMLRABYSGSFRTSNALSATFTTNLTQYSLSRVDLGD-----CVGKEAREADRIYL 449
Qy 395 AKVNSTHSGYDLKYFTKDGGLYLWQPLONRLLDANK- LNNETYSRRSRQAESTTD 453
Db 450 EKNYNTHLRVGSQYLYATGGFLIAYQPLLSNNLADLYVKELMREQAOKPEERKLNATTD 509
Qy 454 PMEMTNGAGGESSENSITVAQVOYADNLRIRINNILEDLSKACROHRAALVWNE 513
Db 510 -----GKVITTSSEVFARLQTYNHIQKHVNEFMGRWAVSWCELOQELLTWN 559
Qy 514 LSKINPTSVMSIYNRPVSARKIGDVISVSNCTIWDQTSVSLHKSRLLSASDEKCFSRP 573

Db 560 AKKINPSAISVTLHRRVSGACMLGDLVLAISTCVAVPAENVIMQNSMRIPS-KPGTCYSRP 618
Qy 574 PVTFRKWNDSITYKGGOLGVNNEILLTTTILETCOENTEYFYFOAKTDWYIYKNVHLKTVP 633
Db 619 LJSFKHVDGEELMEGOLGNNELRLDRDAVEPCVSGHKRYFLFGAGVVFFEETYSHOLS 678
Qy 634 LSSITTLDTFIALNFTLLENVPKVIETLYTRDRSKLSNVFDIETMPREYNYAQRVSGLR 693
Db 679 RSDITAVSFIDNLNITWLEDHFPVLEVTRQEIKDSGLDYAEVORRNQHLARADI- 737
Qy 694 KOLLDLSTNRNQFVDAFSGSLMDLGAQGQTVVNAVSGVATLFSIVTGFNFINKNPGGM 753
Db 738 DTVIKADPNAATFAGLHG-FFEGLDGVDGRAVGWLVGVGVVATVSGVSFLSNPPGAL 796
Qy 754 LMIIVIGVLFALYFLTKTKIYETAPIKMIYPEIDK-LKERE-----GKSEIAPISEEL 808
Db 797 AIGLLVLGLVAFAAFYVYVRLQRPNMKALYPLTTTKDLKHPSEGGGGEAMDFDEQKL 856
Qy 809 ERIVLAMHIHQNSHMETKTRDKPDSILTRAQNMUKRSGVSNLKNABESVEML 862
Db 857 DEARSMIKYMALVSAME-----RYKHKAGRRGGTSAILNARLTDWV 897

RESULT 36
B48474
glycoprotein B - feline herpesvirus 1 (fragment)
C;Species: feline herpesvirus 1
C;Date: 07-Apr-1994 #sequence_revision 18-Nov-1994 #text_change 26-Aug-1999
C;Accession: B48474
R;Spatz, S.J.; Maes, R.K.
Virology 197, 125-136, 1993
A;Title: Immunological characterization of the feline herpesvirus-1 glycoprotein B and
A;Reference number: A48474; MUID:94025559; PMID:8212548
A;Contents: C-27
A;Accession: B48474
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-943 <SPA>
A;Cross-references: GB:S6371; NID:g435792; PIDN:AAB28559.1; PID:g435794
A;Note: sequence extracted from NCBI backbone (NCBIN:138805, NCBIPI:138807)
C;Superfamily: herpesvirus glycoprotein B
C;Keywords: glycoprotein

Query Match 17.4%; Score 789.5; DB 2; Length 943;
Best Local Similarity 27.3%; Pred. No. 5.7e-39;
Matches 231; Conservative 151; Mismatches 364; Indels 101; Gaps 27;

Qy 80 PSTFPYRVCSAGVGDFRFTQDHVCPDAS-DMVHSEGLIYKQNIIPMFVRVRYKV 138
Db 121 PSTF--YMCPPSPSGSTVVRLEPPRACPDYKLGKNFTGEGIAVIFKENIAPYKFKANIYKN 178
Qy 139 VTTSTVYNGIYSDITNQHTFYKSIEPWE-TEKMDTIYOCFNSRLNLTGNNLLTYVDRDD 197
Db 179 IIMTTVMWSSGYAVTTRNRYTDRVPVKVQBITDLIDRRGMCLSKADYVVRNYYQTFADRDE 238
Qy 198 INMTVFLQVDGVTPDVVKRYGQPELYLEPCWFWS-----YRRRTTVNCELMDM 247
Db 239 DPRELPLKPPSTLSRVR-----GWHNTNITYTKVILLDFHSGTSCVNCIVEEV 286
Qy 248 FARNSNPPDFVTATGDTVEMSPFWSGED-----DHENKMKHEKPFVSVINNYKVVDYQNR 303
Db 287 DARSVYVYDSFALSTGSDVIHMSPPFFGLRGDAHVEHTSYSSDR---FQIEGYYPIDLDTD 343
Qy 304 GT-VPLGKTRIFLDREBYTLSEKHLKNMSYCPILTWKAFYNGIQTEHSGSYHFVANDIT 362
Db 344 YTGAPV--SRNLETPHVTVAWNTPKSGRVCTLAKWREI-DEMLPMNITGYSYRFTAKTIS 400
Qy 363 ASFTTSKEDMKENITY--HCLNEEIKAEIKKY-AKVNSTHSGYDLKYFTKDGGLYL 419
Db 401 AFISNTSQF-EINRIRLGDCAKAEADRIYKSKYKTHIQTGTLTYLARGGFLIA 459
Qy 420 WOPLIQNRL-----LDANKLN--NETYSSRSR-----ROAESTTDP 454

```

Db 460 FRPMISNELAKUYINELARNSRNTVVDLSALLNPSGETVQTRRSVPSNQHRRSRSTIEG 519
Qy 455 MMEMTGAGGSEYSENSITVAOVAYDNLAIRINNIILEDLSKAWCREOHAALVWNEEL 514
Db 520 GIETVNNAS--LLKTTSSVEPAMQPAYDYIOAHVNEMLSRITATWCTLQNHVLTWTET 577
Qy 515 SKINTPSVSMIYNRPVSAKRIGDVISNSCIVDPQTSVSLHKSURLLSASDEKCFSRPP 574
Db 578 LKLNPGGVVSMALERRVSARLLGDVAVTCQVNISSGHVYIQNSMR-VTGSSTTCYSRPL 636
Qy 575 VTEKFWNDSTIYKGLGVNNEILLTTTLYLETQENTYEFQAKTDMYIKYKNEHLKTVPL 634
Db 637 VSFRLANDSEXTEGOLGENNELVERKLEPCTVANKRYFKFGADYVVFEDYAYYRKVPL 696
Qy 635 SSITTLDTFIALNFTLLENVDFKVIELYTRDEKRLSNVFDIETMREYNYAORVSGLRK 694
Db 697 SEIELISAVY-IKSTILLEDEF-LHSSYTRALEBTGDPDYSEIORRQLHALKFYDI-- 752
Qy 695 DLLDLSTNRNQVDAFGLMDLGAAGVQTVVNAVSVATLFSIIVTGFINFKNPFPGML 754
Db 753 DSIVRVDDNVLVIMRGWANPFOGLGDVGAGFGKVVILGAASAVISTVGVSSFLNPFGLA 812
Qy 755 MIIVVIGVLPALYFLTKTKIYETAPIKMIYEPIDKLKERCKSE-----IAPI 803
Db 813 VGLLLAGIVAFLAYRYSRLRANPMKALYPTVTRNLKQAKSPASTAGSDSPGVDDF 872
Qy 804 SEELER-----IVLAMHIHOONSHMETKTRKDPKDSILT-RAONMLRKRS--YS 851
Db 873 DEELMQAREMIKYMSLVSM---EQECHKAMKNGP--AILTSHLTNMLARRGPKYQ 927
Qy 852 NLKNAES 858
Db 928 RLNNLDS 934

RESULT 37
JH0109
glycoprotein 14 precursor - equine herpesvirus 1
C:Species: equine herpesvirus 1
C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 26-Aug-1999
C:Accession: JH0109
R:Guo, P.
Gene 87, 249-255, 1990
A:Title: Characterization of the gene and an antigenic determinant of equine herpesvirus
A:Reference number: JH0109; MUID:90236317; PMID:1692002
A:Accession: JH0109
A:Molecule type: DNA
A:Residues: 1-979 <GUO>
A:Cross-references: GB:M34861; NID:g330900; PIDN:AAA46086.1; PID:g330901
A>Note: glycoprotein 14 is homologous to corresponding glycoprotein of pseudorabies, bo
C:Genetics:
A:Gene: gp14
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein; transmembrane protein
F:1-86/Domain: signal sequence #status predicted <SIG>
F:87-979/Product: glycoprotein 14 #status predicted <MAT>
F:831-871/Region: hydrophobic
F:165,275,380,423,497,514,560,727,749,951,970/Binding site: carbohydrate (Asn) (covalent
Query Match 17.1%; Score 777.5; DB 2; Length 979;
Best Local Similarity 26.3%; Pred. No. 3.1e-38;
Matches 247; Conservative 159; Mismatches 377; Indels 155; Gaps 34;
Qy 32 VTSPNWTATWSTESPLTHGYTHDSSHGEGNENRDSEONKNYSGSPFTFPRVC-SA 90
Db 83 VRAVPTTSPPTSTETS--MSTH--SHG-----TVDPTLLPTETDPLRLAVRE 127
Qy 91 SGV-----GD-----VRFQTDHVCDDAS-DMVHSGILLIYKQNIIPFMRVYRK 134
Db 128 SGILEADGDFYTCPPPTGTVRIEPPPTCPKDFGRNFTGIAVIFKNIAPYKFRANV 187
Qy 135 YRKVVTSTTVNGIYSDSITNQHTEYKSGIEPWETEK---MDTIYQCFNS---LRLNTGG 187

```

```

Db 188 YYKDIVTVRVWVGYSHTSLSDR---YNDRVPSVVEEIFGLIDSKGCKSSKAEYLRL---D 240
Qy 188 NLLTYV---DRDDINMTVFLQPDGVDPDKRYGSQELYLEPQWF-WGSYRRRTTNVCE 243
Db 241 NIMHAYHDDDEVELD--LVPSKPAFGARAWOTNTDTTSYVGWMPWRHY-TTSSVNCI 297
Qy 244 LMDFAFSRNPDPDFVTATGTVMSFPW----SCEDDHENKMKHEKPMFVSVINNYKVD 299
Db 298 VEEVEARSVPYDVSFALSTGDIVVASPFYGLRAAAIEHNSVAQER---PRQVEGYRRPD 354
Qy 300 YONRGTVPLGKTRIFLDRBEVTLNWEKHLKMSVCPILTKWAFYNGIQTGSHSGSVHFVAN 359
Db 355 LDSLQAEPEPVTKNFIPTPHVTVSNWTEKKEVEACTLTWKKEVDELVRDEFRGSYRFTIR 414
Qy 360 DITASFTTSKEDMK-EFNTTVHCLNEEIKAEIKKAK-VNSTHSHKYGLDKYFKTDGGLY 417
Db 415 SISTSIFNTTQFKESAPLTECVSKAKEAIDSIYKKQYESTHVFSGDVEYIYLRGGFL 474
Qy 418 LVMOPLI-----QNRLLDAKNKL-----NNETYSRRSRROAESTTTPMMEMT 459
Db 475 IAPRMLSNELARLYLNLVRSNRTYDLKNLLNPNANNNTTTRRSLLSVPEPQPTQD 534
Qy 460 G-----NGAGGEYSSE-----NSITVAQVQVAYVNLRIIRINI 492
Db 535 GVHREQILHRLHRAVEATAGTDSNVTAQKLELIKTTSSIEFAMLPQAYDHIQSHVNEM 594
Qy 493 LEDLSKACRQHRAALVWNELSKINPTSVMSIYNRPVSAKRIGDVLVSNCIVVDQTS 552
Db 595 LSRIATACWPLQNKERPLNEMVKITPSAIVSATLDERVAARVLGVDVIAITHCAKI-EGN 653
Qy 553 VSLHKLRLLSASDEKCFSRPPTF---KFMNDSTIYKGLGVNNEILLTTTLYLETQEN 609
Db 654 VYLQNSMR--SMDSNTCYSRPPTVFTITQANNRSGTEGQCEENEIFETERKLEPCALN 711
Qy 610 TEYFQAKTDMYIKYKNEHLKTVPLSSITLDTFIALNFTLLENVDFKVIELYTRDEKRL 669
Db 712 QKRYFKGKEYVYVYENYTVFRKVPPTIEVISTVVELNLTLEDREFLPLEVYTRAELED 771
Qy 670 SNVFDIETMREYNYAQRVSLRKDLLDLSTNRNQVD-----AFGSLMDDLGA 719
Db 772 TGLLDYSEIQRRNLHALR-----FYDIDSVVN--VDNTAVIIRGSPAFSAWVKWR 822
Qy 720 VQOTVNAVSVATLFSIIVTGFINFKNPGGMLIIVIGVLPALYFLTKTKIVETA 779
Db 823 PWERSFARGAVV---STVSGIACFLNPFGLAIGLLVIAGLVAFAFFAYRYMQIRSN 878
Qy 780 PIKMIYPEIDKLKERCKSEIAPISE-----EELERIVLAMHIHQONSHMETKT 828
Db 879 PMKALYPITTKALKNKAKTSYQONEEDGSDFDDEAKLEAEAREMIKYMSVMSVALEKQEKKA 938
Qy 829 RKDPKDS-----ILTRAQNMLRKRS--YSNLKNAESVE 860
Db 939 IK--KNSGVGLIASNKLALRRRGPKYTRLQOQNDTME 974

RESULT 38
VGBEPS
glycoprotein gII precursor - suid herpesvirus 1
C:Species: suid herpesvirus 1
C>Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 16-Jul-1999
C:Accession: A29159
R:Robbins, A.K.; Dorney, D.J.; Wathen, M.W.; Whealy, M.E.; Gold, C.; Watson, R.J.; Holla
J. Virol. 61, 2691-2701, 1987
A:Title: The pseudorabies virus gII gene is closely related to the gB glycoprotein gene
A:Reference number: A29159; MUID:87284141; PMID:3039163
A:Accession: A29159
A:Molecule type: DNA
A:Residues: 1-913 <ROB>
A:Cross-references: GB:M17321; NID:g334053; PIDN:AAA47465.1; PID:g334054
A>Note: the authors translated the codon GAC for residue 860 as Asn
C:Superfamily: herpesvirus glycoprotein B
C:Keywords: glycoprotein; transmembrane protein
F:1-40/Domain: signal sequence #status predicted <SIG>

```

F;751-819/Domain: transmembrane #status predicted <TM>
F;820-913/Domain: intracellular #status predicted <INT>
F;151,261,441,516,633,697/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 17.1%; Score 775.5; DB 1; Length 913;
Best Local Similarity 27.0%; Pred. No. 3.7e-38;
Matches 235; Conservative 165; Mismatches 388; Indels 81; Gaps 26;

QY 26 SIATGVTSPNTATWTEPLCHYTHDSHGCGNENNRSEEQNKNIYSPSTFPY 85
DB 85 SLEIEAFSP-----GPSEAP-DGEYGLDARTAVRAAATERDR-----F 123
QY 86 RVCSASGVGVDFQTDHVCDDASD-MVHSEGIILYKONIIPFMRVRYKRVVTTSTV 144
DB 124 YVCPPPSGSTVRLPEPQACPEVSGQGNFTTEGTAFLVKENIAHKFAHYKYNVITTV 193
QY 145 YNGIYSDSIHQHTFYKSIPEWE-TEKMDTIYQCFNSRLNTGNNLLTYVDRDDINMTVF 203
DB 184 WSGSTYAAITNRFTDRVPVPVQEIITDVIDRRGKCVSKAEYVRNNHKVTAFTDRDENPVEVD 243
QY 204 LQPDVGVTDPVKYSGOPELYLEPGFWGYSRRRTTVNCELMDFMARSNPFFVFTATG 263
DB 244 LRPSRLNALGTRGHHHTNDTYTIG-AAGFYHTGTSVNCIIVEVEARSVYPYDSFALSTG 302
QY 264 DTVMSPFWS-GEDDHENKHEKXPWFVSVNNYKVDYQNRGTVPLGKTRIFLDREBYTL 322
DB 303 DIVMSPFYGLREGAHEHGYAPGRFOQVEHYPIIDLRLRASESVTRNFRTHFTV 362
QY 323 SWEKHLNMSYCPPLLWKAFYNGIQTE-HSGSYHFVANDITASFTT--SKEDMKEFNTRY 379
DB 363 AWDWAPTRRVCSLAKWREAEEMTRDETROGSRFTSRALGFSVSDVTQLDLQRVHLG- 421
QY 380 HCLNEEKABIEKKY-AKVNSTHSGYDGL-KYFKTDGGLYLWQPLIQNLRLDA-KNKLN 436
DB 422 DCVLRASEAIDATYRRRYNSTHVLADRPVVLARGGVVAFPRPLISNELAQLYARELE 481
QY 437 N-----ETYSRRSRQ--AESTDPMEMTNGAGGEVSSSENSITVAQOVAYD 483
DB 482 RLGLAGVVGPAAPAAARRARRSPGAGTPEP---PAVNGTGHRLRITTGSAEFARLQFTYD 538
QY 484 NLRIINNILEDLSKACREQHRALVWNLKINPTSVMSMIYNNRPVSARKIGDVISVS 543
DB 539 HIQAHVNDMLGRIAAAWCELQNKDRTLWSEMSRLNPSAVATAALQQRVSARMGLGDVMAIS 598
QY 544 NCIVVDOTSVLHKSLLSASDEKCFSPRPVTFKFMNDSTIYKQGLGVNNEILLTYYL 603
DB 599 RCVEV-RGGVYVQNSMR-VFGEKTCYSRPLVTFEH-NGTGVIEGQLGDDNELLISRDLI 655
QY 604 ETQCENTEYVFOAKTDMVIYKVEHLKTVPLSSITTLDTFIALNFTLENVDFKVIELYT 663
DB 656 EPTGNNHRRYFKLGSVYVYEDYVYVRNVEPE--TISTRTVNLTLLEDFLEPLUEVYT 713
QY 664 RDEKRLSNVFDIETMFREYNYYA-----QRVSLGRKDLDDLSTNRNQVDAFGSLMDDL 717
DB 714 REELADTGLLDYSEIQRRNQLHALKFYDIDRVKVDHNVLLRGIANFF-----QGL 765
QY 718 GAVGQTVNAVSGVATLFLSSIVTFINFIKNPFGMLMIIVGLVPAIYFLTKYIYE 777
DB 766 GDVGAAYGVKVLGATGAVISAVGMVSVFLNPGALAIGLLVLAGLVAAPFLAYRHISRLR 825
QY 778 TAPTKMIIPEIDKLKREGKSEIAPISEBELER-----IVLAMHIIHQONSHMETKT 828
DB 826 RNPWKALYPVTTKLEKGGVDE-GDVDEAKLDQARDMIRYMSIVSAL---EQEHLKARKK 881
QY 829 RKDPKDSILTRAQNMRLKRSGYSNLKNAE 857
DB 882 NSGPA-LLASRGVAMATERRHYQFLESED 909

RESULT 39
VGBEEG
glycoprotein g1 precursor - bovine herpesvirus 1
N;Alternate names: glycoprotein 11a; glycoprotein 16; glycoprotein g130; glycoprotein gV

C;Species: bovine herpesvirus 1
C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 16-Jul-1999
C;Accession: A31166
R;Mista, V.; Nelson, R.; Smith, M.
Virology 166, 542-549, 1988
A;Title: Sequence of a bovine herpesvirus type-1 glycoprotein gene that is homologous to
A;Reference number: A31166; MUID:89020821; PMID:2845560
A;Accession: A31166
A;Molecule type: DNA
A;Residues: 1-928 <MS>
A;Cross-references: GB:M23257; NID:G340858; PIDN:AAA46013.1; PID:G511852
A;Note: the authors translated the codon CTG for residue 534 as Ser
C;Superfamily: herpesvirus glycoprotein B
C;Keywords: glycoprotein; transmembrane protein
F;1-67/Domain: signal sequence #status predicted <SIG>
F;68-928/Product: glycoprotein g1 #status predicted <GGI>
F;756-824/Domain: transmembrane #status predicted <TM>
F;105,153,442,484,579,637,703/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 15.9%; Score 722; DB 1; Length 928;
Best Local Similarity 26.2%; Pred. No. 5.7e-35;
Matches 240; Conservative 156; Mismatches 427; Indels 94; Gaps 26;

QY 2 AGSLKLRGSLALWLYLQVALYSLSAETGVT-----SPPNTATWSTESPLTGHY 51
DB 41 AGCARAALAAALLWATWALLLAAPAGRAPATTPAPPPEEASAPAPPSPFGPDGDDA 100
QY 52 GTHDSSHGGRGNENNRDSEQNKNIYGPSPTFPYRVCSASGVGDVFRFQTDHVCDD-ASD 110
DB 101 ASPDNSTDVRAALRLAQAAGENSRRF-----VCPPPSGATVTVRLAPARCPPEYGLG 151
QY 111 MVHSEGIILYKONIIPFMRVRYKRVVTTSTVYNGIYSDSIHQHTFYKSIPEWE-TE 169
DB 152 RNYTEGIGVIYKENTAPYTFKAIYYKYNVITVTWAGSTYAAITNQYTDROPVGMGEITD 211
QY 170 KMDTIYQCFNSRLNTGNNLLTVVDRDDINMTVFIQPDVGTVPDVKRYGSOPELYLEPGW 229
DB 212 LVDKKWRCLSKAEYLRSGKVVAFDRDDDPWEAPLKPALSAFGVGRWHTDDVYTAGL- 270
QY 230 FWGSVRRRTTVNCELMDFMARSNPFPDFVTATGDTVMSPFWS-GEDDHENKMKHEKWP 288
DB 271 SAGLYRTGTSVNCIIVEVEARSVYPYDSFALSTGDIYYMSPPFYGLREGAHRHTSYSPER 330
QY 289 VSVNNYKVDYQNRGTVPLGKTRIFLDREBYTLSEKHLKNMSYCPPLLWKAFYNGIQT 348
DB 331 FOQIEGYKRDWATGRRLKEPVSRLFRTOHVTVAVMDVFKRVKNCVSLAKWREADMLRD 390
QY 349 EHSGSYHFVANDITASFTTSKEDMKEFNTRY-HCLNEEKABIEKKY-AKVNSTHSGYKD 406
DB 391 ESRGNFRFTARLSATFVSDSHTFALQNVPLSDCVIEEAEAAVERVYRERYNGTHVLSGS 450
QY 407 LKYFKTDGGLYLWQPLIQNLRLDAKNKL-----NNETYS-----RRSRQAE 449
DB 451 LETYLARGGVVAFRPMLSNEL-AKLYLQELARSNGTLEGLFAAAAPKPGPRARRPRR 508
QY 450 SITDPMEMTNG-----AGGEYSSSENSITVAQOVAYDNLRIINNILEDLSKAWC 501
DB 509 LRPAPE-----GRQARRRRRHAGGRVTVSLAEFAALQFTHDHTRTSEHHV-HRLASPPC 562
QY 502 REQHRAALVWNLKINPTSVMSMIYNNRPVSARKIG-----DVI--SVSNCIYVVDQTSVSL 555
DB 563 LIQNKERALLWAAAKLNPSAAASALDRPPRACWGTGHRDVLPRAGRALFIENSARA 622
QY 556 HKSLRLLSASDEKCFSPRPVTFKFMNDSTIYKQGLGVNNEILLTYYTLETQENTYFFQ 615
DB 623 RR-RLL-----QPPRFLSFGNESEPEVGOLCEDNELLPGRELVEPCTANHKKRYFR 671
QY 616 AKTDMVIYKVEHLKTVPLSSITTLDTFIALNFTLENVDFKVIELYTDEKELSNVFDI 675
DB 672 FGADYVYVYENAYVRRVPLAELEVISTFVDLNLTVLEDFLEPLVYTRAEADTGLLDY 731
QY 676 ETMFREYNYYAQRVSLGRKDLDDLSTNRNQVDAFGSLMDDLCAVGQTVNAVSGVATLFL 735

Db 732 SEIQNRNQJHURFYDI--DRVVKTDGNMAIMRGLANFFQGLGAVGQAVGTVVLAGAGAA 789

QY 736 SSVITGVFINPKNPFQGMMLMIIWIGVLFALIVPLTKTKTIYETAPIKMIYPEIDK-LKER 794

Db 790 LSTVSGIASFIANPFQALATGLLVLAGVAAFLAYRISRLSRNPWKALYPITTRALKDD 849

QY 795 EKKSEIAPISBEE-----LER----IVLAMIHQQNSHMETKTRDKPKDS--ILTR-AQN 842

Db 850 PGRNR--PGEEBEEFDAKLEQAREMIKIMSLVSAVERQEHAKKSKNAARLLATRLTLQ 907

QY 843 MLRKES--CYSNLKNAE 857

Db 908 ALRRRAPPEYQQLPWAD 924

RESULT 40

S68553

surface layer protein tetraabrachion precursor - Staphylothermus marinus

C:Species: Staphylothermus marinus

C>Date: 27-Nov-1997 #sequence_revision 12-Dec-1997 #text_change 17-Mar-1999

C:Accession: S68553; S78081

R:Peters, J.; Baumeister, W.; Lupas, A.

J. Mol. Biol. 257, 1031-1041, 1996

A:Title: Hyperthermostable surface layer protein tetraabrachion from the archaeobacterium cure.

A:Reference number: S68553; MUID:96192087; PMID:8632466

A:Accession: S68553

A:Molecule type: DNA

A:Residues: 1-1524 <PET1>

A:Cross-references: EMBL:U57967

A:Experimental source: strain F1

A:Accession: S78081

A:Molecule type: protein

A:Residues: 40-45;91-95;151-158;327-331;445-449;490-497;499-505;596-614;617-624;638-648;5-1451 <PET2>

A:Experimental source: strain F1

C:Complex: heterotrimer; two alpha chains and two beta chains

C:Keywords: blocked amino end; glycoprotein; membrane bound

F:1-26/Domain: signal sequence #status predicted <SIG>

F:27-700/Product: surface layer protein tetraabrachion alpha chain #status predicted <MAT>

F:701-1524/Product: surface layer protein tetraabrachion beta chain #status experimental

F:1019-1029/Region: hinge

F:1495-1519/Domain: membrane anchor #status predicted <MBN>

F:144,605,641,685,708,1279,1402/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 3.6%; Score 162; DB 2; Length 1524;

Best Local Similarity 18.1%; Pred. No. 0.22;

Matches 148; Conservative 132; Mismatches 301; Indels 236; Gaps 35;

QY 24 SLISIAETGVTSPPNTATWSTESPLTCHYGTGTHDSHGERGNENRDSERQN-----KNYIGS 79

Db 729 TLSI--TVVGAAPN-----TEYNFTFGYQVHDLNIGITRISYSPQWGTWISLVTDIYGT 781

QY 80 PSTPYRVCSAGVGDFRFQDHPDASDMVHSEGIILYKQNIQIPMFVRVRYKRW 139

Db 782 GST-----SVPLITLYPTSIVINATMDVI---TWLRSGSGTDLDFSV----- 822

QY 140 TTSTVYNGIYSDSITNQHTFYKSIPEWTERKMDTIYQCFSRLNTGSL-----LTYV 193

Db 823 --DVSYNG-FIDNLTPITYVFG-----PSDTPGSGFN-IYVNTTYNVSVVRVAVDYL 871

QY 194 DRDDINMTV--FLQPDVGVTDPVKRYGQPELYLEP-----GWFGWSYRRRTTVN 241

Db 872 PRTNVVISPEVTLPGDITVQIFPHHNEVWGFIETALFDENQLLGY-----LT 922

QY 242 CELMDMFARSNPPDFPFVATGDTVE--MSPFWSGE---DDHENKMKHPWFV----- 289

Db 923 VRLVDPLSNT-----VVERVAGYVAGNLIVEDVDGDGNEVFWVNLTAFLV 969

QY 290 -SVINNYKV-----VDYQNRGTVPGLKTRIFLDREYTLSEKHLKNNMSYCPFLTWKAFY 343

Db 970 LGVDKTYRVDVELFLAVLNPSSNITGVTA--DNECYV---QLDLNGTIY-----WNLGL 1019

QY 344 NGIQTEHSGSYHFVANDITAGFTTTSKDMKBFNNTYHCLNEEIKAEIEKKYAKVNSTHSHK 403

Db 1020 SGIMLGGDQIVTVLVGLEKLDITKDGIAEINATVNDINTYLVKNVTDLLKTINNS--- 1076

QY 404 YGDLKYFTKDGGLYLVWQPLIQNRLLDAKNKLN--NETYS-----RRSRQA 448

Db 1077 ---VVMKNDTATLIIGQAEIKAKLDLNLTSQVNDTVTMLACCNNAKSVLNRMEGTL 1133

QY 449 ESTTDPNMEMTGAGGEYSSENSITVAQVQYVDNLRIRINNILEDLSEKAWCREQHRAA 508

Db 1134 NSTYGVNLVKSD--LSTLIDTVNNVPIPKFNELVNVTVEINASRD----- 1178

QY 509 LVNNELSKINP--TSVMSMIYNRPVSARKRIGDVISVNCIVVDQTSVSLHKSRLLSASD 566

Db 1179 LIIQKISSVNDSLTIIISAGFN-----DVEAMISNLTLLNLRID 1218

QY 567 EKCFSRPVTPKFWNDSTIYKQGVNNEILLTTTLETQCENTEYFYFOAKTDMYIK-- 624

Db 1219 E-----LEGTL-----LFYMTANEQRLEGIINETADDIVYRLT 1251

QY 625 ----NYEHLKTVPLSSITLDTET-----ALNETLLENVDFKVIETREDEK 667

Db 1252 VIIDRYESLKNLITLRADRLMIINDNVSTILASIGNVNLTVPFNKLNDEIELGDNAT 1311

QY 668 RLSNVFDIETMFREYNYVAQRVSLRKLDDLSTNRNQFVDAGSLMDDLGAVGOTVVNA 727

Db 1312 INAGIFQISNLGNANQ-----LLDLTSSKVELLNAISSNASSAISSEIHNA 1359

QY 728 VSGVATLPSSI-----VTG-----FINFIKNPFGGM 753

Db 1360 VNQLSTVLVQVNDTLTKITGEADNINLFLSSLEGSM 1396

RESULT 41

T28238

ORF MSV077 hypothetical protein - Melanoplus sanguinipes entomopoxvirus

C:Species: Melanoplus sanguinipes entomopoxvirus

C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000

C:Accession: T28238

R:Afonso, C.L.; Tulman, E.R.; Lu, Z.; Oma, E.; Kutish, G.F.; Rock, D.L.

J. Virol. 73, 533-552, 1999

A:Title: The genome of Melanoplus sanguinipes entomopoxvirus.

A:Reference number: 220484; MUID:99102612; PMID:9847359

A:Accession: T28238

A>Status: preliminary; translated from GB/EMBL/DBD

A:Molecule type: DNA

A:Residues: 1-598 <AFO>

A:Cross-references: EMBL:AF063866; NID:g4049647; PIDN:AAC97632.1; PID:g4049672

C:Genetics:

A>Note: MSV077

Query Match 3.5%; Score 158.5; DB 2; Length 598;

Best Local Similarity 21.1%; Pred. No. 0.087;

Matches 106; Conservative 88; Mismatches 192; Indels 117; Gaps 28;

QY 318 BEYTLSEKHLKNNSY-----CPLTLWKAFFNGIQTEHSGSYHFVANDITASFTTSKEDMK 373

Db 169 DEYSNIFKRHCNSLKYIIGKILV-----NDLQVK---EYIWDN-----SHKFDIK 212

QY 374 ---EFNTTY-HCLNEEIKAEIEKKYAKVNSTHSHKYGDLKY-----FKTDGGL 416

Db 213 FSEQYFITPIFYVNTGVLTGFKFENGFAELYSISSNQIIKYDHVITCKSDFINIENG- 271

QY 417 YLVWQPLIQNR-LIDAKNKLN-NETYSRRRRRQAESTTDPMENTGAGGEYSSENSIT 474

Db 272 YFISMPFNEGQTVLSGRELNVWSQTFGKTCIRNTENKHFVLNVTLANSLKNEY-----IN 326

QY 475 VAQVQYAYD-NLRIRINNILEDLSEKAWCREQHRAALVWNELSKINPTSVMSMIYNR-PVS 532

Db 327 IARFQEVLDKKLKIITKNL-----DTLCKNYVIYRKIYTEICKNNPQCMKYFLNHYNVK 381

QY 533 AKRI-GDVISVSNICVVDDQTSVSLHKSRLLSASDEKCFSRPPVTFK-----FNWDS 583

382 ARYGGNIGIKYCIETDFEI-----IKOYNKTDKCEIFVPMKINNNNTKFGYMNPS 436
QY 584 TIYKQGLGVNNEILLTTTLETQENTYFQAQTDMMYIKYVYEHKLTVPSSII--TTLD 641
Db 437 T-----NEVFTSPKNDYCSIT-YVDINDLMIYISNNQINVSQKIHTYGHETLL 486
QY 642 TFIALNFTLNVDFKVIETDRKRLSNVDFIETMFREYNYAQRVSGRLKDLLOLST 701
Db 487 TNKISPIILKVIN--ISDLYF-BQSTFQDVYDI-----AYHFEDELINPKQTYNFTS- 537
QY 702 NRNOFVDAFGLMDLGLAVGQTVVNAVSGVATLFSSTIVTGFINFKPFGMLMIIVIG 761
Db 538 -----SMDPLNFHG-----IPIHIFIGFVIF-----GIFLVRYS 569
QY 762 VLFA-IYFLTKKTK-IYETAPIK 782
Db 570 QIFNSCASIFRKTIVYENIPMK 592

RESULT 42
T28677
rhoptry protein - Plasmodium yoelii
C;Species: Plasmodium yoelii
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 01-Dec-2000
C;Accession: T28677; C45521
R;Keen, J.; Sinha, K.; Brown, K.; Holder, A.
Mol. Biochem. Parasitol. 65, 171-177, 1994
A;Title: A gene coding for a high molecular mass rhoptry protein of Plasmodium yoelii.
A;Reference number: Z20508; MUID:95021522; PMID:7935623
A;Accession: T28677
A;Status: preliminary
A;Molecule type: DNA
A;Molecule type: translated from GB/M34283
A;Residues: 1-2269 <KEE>
A;Cross-references: EMBL:L27838; NID:9457145; PID:9457146; PIDN:AAA21304.1
R;Keen, J.; Holder, A.; Playfair, J.; Lockyer, M.; Lewis, A.
Mol. Biochem. Parasitol. 42, 241-246, 1990
A;Title: Identification of the gene for a Plasmodium yoelii rhoptry protein. Multiple clones.
A;Reference number: A45521; MUID:91101660; PMID:2270106
A;Accession: C45521
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 2131-2269 <KE2>
A;Cross-references: GB:M34283

Query Match 3.4%; Score 154; DB 2; Length 2269;
Best Local Similarity 18.0%; Pred. No. 1.2;
Matches 151; Conservative 154; Mismatches 306; Indels 228; Gaps 41;

QY 93 VGDVFRFQTDHVCDDASDMVHSEGILLIYQNIIPFMRVRKRVKVTSTVY--NGIYS 150
Db 91 IRELFFKYLD-----EEAEKYLEGLKLELNKKIKDIIAKIEYKNTVELKKEKNWAYI 146
QY 151 DSITNQ-----HTFYKSIETPWEKMDTIYOCFNSRLNTGGLNLTGVDRDDI 198
Db 147 DELANQSPYKVTGYENKNTIYNTIKSY---PQIYE-----GDIITFY--NEL 190
QY 199 NMTVFLQPDVQVTPDVR-----YCSQPLEYLPFGWGSYRRRTVNCCLMDMF 248
Db 191 SSVIKEDPIDDIEDKTKLENLRSKIDNVYDKIQKMEIE-----TVKSHLNIE 238
QY 249 ARSNPPDFFTATGDTVENSFPWSGE-----DDHENKMKHKEKPFVSVNNYKVV 299
Db 239 TNNKLP-----NTILEIKYIYDEISKELNKMLEDFFKNKE-----LSN-KISD 282
QY 300 YQNRGTVPGLKTRIFLDREYTLSEWHLKMSYCPPLTLWKAFYN-----GIQTEHSGS 353
Db 283 YDKK-----REQLSYKSKMELRNH-----YNSQTVNDTKKEEBAQ 320
QY 354 YHFVANDITASFTSKEDMKFNFTYHCLNEIKAEI-----EKKYAK-VNSTHSGYD 406
Db 321 NYDKSNHMTTIPNEDEISKISEVTKMKDEILSKYNTYIDFNKKYKETYNSHESQFTE 380
QY 407 LKVFKTGGGLYLVWQPLIQNRLLDAKNKLNNETYSRRSRQAESTTD--PMMEWTGNAG 464

Db 381 L-----ID-----KIKAEVSDRELKKCEQSFNDKNSLINETKNSIE 416
QY 465 GEYSSENSITVAQVAYDNLRI--NNILEDLSKAWCREHQRAALVWNELSKINPTSV 523
Db 417 KEYQNINTLKKVD-EY-----IKVCKSTKESITKFSKQTIKMDLNLQNIKVTKETNSI 469
QY 524 SMYLRPVSAKRIGDVISVSNICVVDOTSVLSHKSRLLSASDEKCFSPRPVTFKMNDS 583
Db 470 DKYIEKFEQILTKQTKLENKF-----TEFSLNHH-----BANNNE-----LIKFS 513
QY 584 TIYKQGLGVNNEILLTTTLETQENTYFQAQTDMMYIKYVYEHKLTVPSSII--TTLD 641
Db 514 --LKANLGINEENLNYQFTE--KEKT--FNDIKEKNIHNEISKIEIKIHASIYNIS 566
QY 636 SITTLDTFIALNFTLLENVDFKVIETDRKRLSNVDFIETMFREYNYAQRVSGRLK 695
Db 567 EETERE--IGIN--IESLTKVFE--KVKNVTNLNKIKEKLKHY-----D 606
QY 696 LLDLSTNRN-QFVDAFGSLMDDLGAQVTVVNAVSGVATL-----PSSIVTGFINFIKN 748
Db 607 FSDFGKEGNIKYTDKIKKINDDIMAVSQIIDQHINGLDDIQKXSESIVSEMKGQINKLEK 666
QY 749 PFGGMLMIIVVIGLFAIYFLTKTKIYETA--PIKMIYPEIDKLKEREKGSBIAPISEE 806
Db 667 VSNTEISNDNVEGI-----KKQOIIVTKIDKKNYIEEINKLSEISKIEKDNTSLE 719
QY 807 ELERIVLAMIHOONSHMETTKTRKDPKDSILTBAQNLKRSGY----SNL-KNAESVE 860
Db 720 KVKDINLSYQNLGNLFLEQIDBEKKK-----AENTIKSMEAYIDDLNKKKSQIE 772

RESULT 43
F82885
hypothetical protein UU482 [imported] - Ureaplasma urealyticum
C;Species: Ureaplasma urealyticum
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C;Accession: F82885
R;Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.
submitted to GenBank, February 2000
A;Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a min
A;Reference number: A82870
A;Accession: F82885
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-4688 <GLA>
A;Cross-references: GB:AE002145; GB:AF222894; NID:G6899476; PIDN:AAF30894.1; GSPDB:GN0012
A;Experimental source: serovar 3; biovar 1
A;Genetic code: SGC3

Query Match 3.2%; Score 146; DB 2; Length 4688;
Best Local Similarity 19.3%; Pred. No. 10;
Matches 184; Conservative 120; Mismatches 349; Indels 300; Gaps 42;

QY 92 GVGDFRFTQDTHVCDDASDMVHSE-----GILLIYQNIIPFMRVRKRVKVT--TSTV 144
Db 2337 GVGNIETIQDRDVHNLNSAKIRFELDLNVLNSDEQATITNNNTTSKAVITDQNKY 2396
QY 145 YNGIYSDSITNOHTFYKSIET---PWETEK-----MDTIYQCFN----- 179
Db 2397 LEATFNLVLNKDITINKIEFNTKPKNASKNGICINTNVIYDATNLIINNDLKITGLPHT 2456
QY 180 -----SURLNTGGNL-----LTYVDRDDIN--MTVFLQPDVQVTPDKRYGSG 220
Db 2457 LKBEFANNKNTISVSLDITNNHISKNLYFIKAFDSNDGQSVLTNPFIADKIVTNNNNK 2516
QY 221 PEL-----YLEPGWFGSYRRRTVNCCLMDMFARSNPPDFFTATGTVEM 268
Db 2517 KETFLNLTLNLSNRQYIFKGLYVYSSNQTNID--ENNKFKNSNVYKIVTKPTITTIQK 2575
QY 269 SPWSCEDDHENKMKHKEKPFVSVNNYKVVYQNRGTVPGLKTRIFLDREYTLSEWHLK 328

Db 2576 NGNWTFPQNAQK-----FKFNINSNDVDF-----STDLDAITFSDQHQTKTITTKL 2625
Qy 329 K---NMSYCPPLTLWKAIFYNG-----IQTEHSGSHFVANDITASF 365
Db 2626 KQKDNQWYEDTISLDLAYNDTYKLLDISITKPMNFAFKIQIE-----NKEQISF 2676
Qy 366 TTSKEDMKFNNTTYHCL--NEETKABIEKKY-AKVNSTHSGYKGLKYFKTGDGLYLVWQP 422
Db 2677 TTQSGPTQLVSIISHDLMNQDNASNOQTITAKVSGVNDLY----- 2718
Qy 423 LIONRLLDKAKLNNTYGRSRQRAESTTDPMMETNGAGGEYSSENSITVAQVQYAY 482
Db 2719 --NNR---KILVVEYRSNNOKSLVESN-----ELTLQDQTOYIFTLPISVANRQYSF 2767
Qy 483 DNLIRIRINLELDSKAWCREQHRALVWNLKINPTSVMSIYNRPVSAKRIGDVISV 542
Db 2768 KEIKITSN-----NNFETLNNSTNVNSFN--VNASKTOIVDN 2805
Qy 543 SMCIVVDQTSVSLHSLRLLSASDEKCFRPPVTFKFMNDSTIYKGLGWNNEILLTTY 602
Db 2806 TN-LATNITSTATITTYKLKSDHVFQVGSIIITYLKSNDDE-----RQEISYTKTI 2856
Qy 603 LETQCENTEYFOAKTDMYIK--NYEHLKT---VPLSSITTL-----DTFIAL 646
Db 2857 TSISDGEATVSFNTDQTLKBEANYKLIKVGFRKPTILAYTNINNDANNVIFEDNNSNY 2916
Qy 647 NF-TLLENVDKVIELYTRD-----EKRLSNVF-----D 674
Db 2917 NFKTLI--VDHKVTNVSSNDSINTTTQTVNIDIDIGIQTWINKIKQLVYTSNDGEILLSD 2974
Qy 675 IETMFREYNYAQRVSGLRKLLDLSTNRNQFVDAFGSLMDD-----LGAVGQTV 724
Db 2975 QXTLLWANNHYSPFLSNLKH-----NRKYLKEVRIINDNKTSIIFHLKNGADWPI 3027
Qy 725 VNAVSGVATLFSISVTGF-----INFIKNP-----FGMLMI--- 756
Db 3028 VNKTQIS--IGSIIPIARAKNNLQSTQIRFILNDPDVNLSEBEAIIYNGENLSVKAK 3085
Qy 757 IVVIG-----VLFAYFLPKTKIYETAPI---KMIY---PIDKUK 792
Db 3086 KVIQSGKYLETFNNLALNQDTIINSITFINKPKIATNIGIDNSNKIYNNNTNINPLK 3145
Qy 793 EREGKSEIAPISEEELERIVLAWHIHOONSHMETKTRDKP--DSLTRAQNM 843
Db 3146 IDNFSVIGVQSDTNKRTL-----EYNNKTNVKNLEKANDNL---QNL 3190
RESULT 44
A38539
p101 protein precursor - Mycoplasma hyorhinis
C:Species: Mycoplasma hyorhinis
C>Date: 31-Jul-1991 #sequence_revision 31-Jul-1991 #text_change 07-Dec-1999
C:Accession: A38539
R.Yogev, D.; Watson-McKown, R.; McIntosh, M.A.; Wise, K.S.
J. Bacteriol. 173, 2035-2044, 1991
A:Title: Sequence and TnphoA analysis of a Mycoplasma hyorhinis protein with membrane ex
A:Reference number: A38539; MUID:91161521; PMID:1848219
A:Accession: A38539
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-888 <YOG>
A:Cross-references: GB:M60787; NID:g150162; PIDN:AAA25422.1; PID:g150163
C:Genetics:
A:Genetic code: SGC3
Query Match 3.2%; Score 145.5; DB 2; Length 888;
Best Local Similarity 19.4%; Pred. No. 0.93;
Matches 163; Conservative 114; Mismatches 293; Indels 269; Gaps 38;
Qy 141 TSTVY--NGIYSDSINQHTFYKSIPEWTEKMDTIYQCFSRLRLMTGNNL----- 189
Db 110 TRVWLLSGLYVDN--QAKTLIKSEPE-EVNLKNIIVKA-----TNAGNQLIFNPMLSAN 161

Qy 190 ---LTVYVDRDDINNTVFLQPV--DGV-TPDVKRYGSOPELYLEBPGFWGYSRRRTTV-- 240
Db 162 EKLKLTFFKDKKNAKETTFEASTNQDGVNIDSEKINLO-----GNYQLVASVIA 210
Qy 241 -----NCELMDM--FARSNPPFPDFVTATGD-----TVMSFPWSGE 275
Db 211 KNNYSLDFNSQVIEWSFPKSNP-----VPTSDVMIKNDKVAIINKINTLELS----- 259
Qy 276 DDHENKMEK-PWFSVSVINNYKV--DYQNGRTVPLGKTRIFLDR-----EETLSWE 325
Db 260 -DHLKPFANIKDFSQLADYEVWLTLMKNQKPAELLAKSTAKIMDKLVFTFDNLTSNTS 318
Qy 356 KHLKMSYCPPLTLWKAIFYNGIQTEHSGSHFVANDITASFTTSKEDMKFNNTTYHCLN-- 383
Db 319 YQISNISY-----KNDYTDKILLDPSINKTFTTNSKPLDFD 354
Qy 384 -BEIKAEIEKKYAKVNSTHSGYKGLKYFKTGDGLYLVWQPLIQNRLLDKAKNKLN----- 436
Db 355 PSNFALKIKDNHYKVNKLSS-----KDNLEAQAQVNVVFTSE 393
Qy 437 NETYSRRSRQRAESTTDPMMETNGAGGEYSSENSITVAQVQYAYD---NLIRINNI 492
Db 394 DKTLEHPAKITQDSTND-----YKIELELDNLQDDQIYSELMIKLAKPTKA 440
Qy 493 LEDLSKAWCREQHRALVWNLKINPTSVMSIYNRPVSAKRIGDVISVNCIIVDQTS 552
Db 441 LKNLGAQDNVIYNAKDDFKISFIN-EKWLNIITLTREVDPKTLTD----- 485
Qy 553 VSLHSLRLLSASDEKCFRPPVTFKFMNDSTIYKGLGWNNEILLTTYLETQCENTEY 612
Db 486 --KHSSYKI-----KLNFSK--IPTVYLNHKALEYQ-SKND-LATTSEVEIKDSSITLD 534
Qy 613 YFOAKTDMYIYKVEHLKTVPLSSITTLDTFIALNFTLLENVDKVIELYTRDEKRSLNV 672
Db 535 YTLDLNLLNVNKEYE-----LKFLIKDNTNNKVIKITKD-LNLNDS 575
Qy 673 FDIETMFREYNYAQRVSGLRKLLDLSTNRNQFVDAFGSLMDDLGAVGQTVVNAVSGVA 732
Db 576 LKILQTFNSFOIKERTINSLNAS--VDINTDLNNSKVDF--SONKQVKVYQEVNNSN--- 630
Qy 733 TLFSSIVTGFNFINKPFPGMLMIIVVIGLFAIVFLTKTKI-----YETAP 780
Db 631 ---THQITGTIKDNNNNKKQL-----DLTFNLELNKKYIKTLEVTNSANDAHNNIV 681
Qy 781 IKMIYPEIDKUKEREGKSEIAPISEEELERIVLAWHIH-----QONSHME 825
Db 682 FNLFTP-----WELQNOFEVSSISDLNKLTKQLDQFQINTNIDLTNSKYKLMKNQNOQV 736
Qy 826 TKTRKDKPKDSLTRAQ-----NMLKRSYGNLKNVSEVEMLN 863
Db 737 KEIVTNQKDSITFNKQNTIISVLUSSIFNQEGKYLIKLLLIANNNNNNFTNAKEISUN 795
RESULT 45
B42771
reticulocyte-binding protein 2 - Plasmodium vivax (fragment)
C:Species: Plasmodium vivax
C>Date: 12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 09-Sep-1997
C:Accession: B42771
R.Galinski, M.R.; Medina, C.C.; Ingravallo, P.; Barnwell, J.W.
Cell 69, 1213-1225, 1992
A:Title: A reticulocyte-binding protein complex of Plasmodium vivax merozoites.
A:Reference number: A42771; MUID:92315338; PMID:1617731
C:Accession: B42771
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1252 <GAL>
A:Cross-references: GB:M80898; NID:g160627; PID:g160628
C:Experimental source: strain Belem, merozoites
C:Genetics:
A:Gene: RBP2

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 15:27:11 ; Search time 26 Seconds
(without alignments)
1564.542 Million cell updates/sec

Title: US-10-055-364-24

Perfect score: 4547

Sequence: 1 MAGSLKRGSLALWLYQV.....KRGYSNLKNAESVEMLNTL 865

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query	ID	Description
1	1949.5	42.9	808	1	VGLB_HSVSA
2	1848.5	40.7	857	1	VGLB_EBV
3	1134	24.9	854	1	VGLB_RHCM6
4	1128	24.8	907	1	VGLB_HCMVT
5	1121.5	24.7	906	1	VGLB_HCMVA
6	1084.5	23.9	822	1	VGLB_HSV7J
7	1055.5	23.2	830	1	VGLB_HSV6U
8	1050.5	23.1	830	1	VGLB_HSV6G
9	1040.5	22.9	830	1	VGLB_HSV62
10	1034	22.7	928	1	VGLB_MCMVS
11	1008	22.2	944	1	VGLB_HSVT2
12	999	22.0	901	1	VGLB_GPCMV
13	900	19.8	920	1	VGLB_HSVSM
14	876.5	19.3	865	1	VGLB_HSVMD
15	857	18.8	932	1	VGLB_HSVBC
16	849.5	18.7	904	1	VGLB_HSVLK
17	845.5	18.6	903	1	VGLB_HSVLF
18	844	18.6	904	1	VGLB_HSV2H
19	841	18.5	904	1	VGLB_HSV11
20	839	18.5	904	1	VGLB_HSV23
21	836.5	18.4	904	1	VGLB_HSV1P
22	836	18.4	883	1	VGLB_ILTVS
23	835	18.4	980	1	VGLB_HSVB
24	833	18.3	885	1	VGLB_HSV2S
25	832	18.3	868	1	VGLB_VZVD
26	830	18.3	883	1	VGLB_ILTVT
27	829	18.2	873	1	VGLB_ILTV6
28	824	18.1	980	1	VGLB_HSV1
29	824	18.1	980	1	VGLB_HSVEA
30	821	18.1	933	1	VGLB_HSV1
31	819.5	18.0	919	1	VGLB_HSVE4
32	805	17.7	917	1	VGLB_HSVB2
33	777.5	17.1	979	1	VGLB_HSVEL

RESULT 1

ID	VGLB_HSVSA	STANDARD;	PRT;	808 AA.
AC	P24905;			
DT	01-MAR-1992 (Rel. 21, Created)			
DT	01-MAR-1992 (Rel. 21, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Glycoprotein B precursor.			
GN	8 OR KCRF1.			
OS	Herpesvirus saimiri (strain 11).			
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;			
OC	Gammaherpesvirinae; Rhadinovirus.			
OX	NCBI_TaxID=10383;			
RN	(1)			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=90163221; PubMed=2154888;			
RA	Albrecht J.-C., Fleckenstein B.;			
RA	"Structural organization of the conserved gene block of Herpesvirus saimiri coding for DNA polymerase, glycoprotein B, and major DNA binding protein.";			
RT	binding protein.";			
RL	Virology 174:533-542(1990).			
RN	(2)			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92333688; PubMed=1321287;			
RA	Albrecht J.-C., Nicholas J., Biller D., Cameron K.R., Biesinger B., Newman C., Wittmann S., Craxton M.A., Coleman H., Fleckenstein B., Honess R.W.;			
RA	"Primary structure of the herpesvirus saimiri genome.";			
RL	J. Virol. 66:5047-5058(1992).			
CC	-!- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; X64346; CAA45631.1; --			
DR	EMBL; M31122; AAA46164.1; --			
DR	InterPro; IPR000234; Glycoprot B.			
DR	Pfam; PF00606; Glycoprotein B; 1.			
DR	ProDom; PD000693; Glycoprot B; 1.			
KW	Glycoprotein; Transmembrane; Signal.			
FT	SIGNAL 1 17			
FT	CHAIN 18 808			
FT	DOMAIN 18 657			
FT	TRANSMEM 658 673			
FT	TRANSMEM 679 698			
FT	TRANSMEM 702 722			
FT	DOMAIN 723 808			
FT	CARBOHYD 30 30			
FT	CARBOHYD 158 158			
FT	CARBOHYD 239 239			
FT	CARBOHYD 251 251			

P08355 pseudorab
P17471 bovine herp
Q00799 plasmodium
P28742 saccharomyc
P24339 schizosacch
P47437 mycoplasma
P36022 saccharomyc
Q02455 saccharomyc
Q23639 caenorhabdi
P19321 clostridium
Q27171 paramedim
Q04956 plasmodium

ALIGNMENTS

```
FT CARBOHYD 285 285 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 331 331 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 344 344 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 355 355 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 361 361 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 532 532 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 569 569 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 587 587 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 598 598 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 808 AA; 91694 MW; FA1BCB89B9D5672E CRC64;

Query Match 42.9%; Score 1949.5; DB 1; Length 808;
Best Local Similarity 46.1%; Pred. No. 3.7e-112;
Matches 373; Conservative 159; Mismatches 236; Indels 33; Gaps 10;

Qy 71 EONKNIYGSPTFYRVCSASGVGVFRFQTDHVCPCDASDMVHSEGLLIYKQNIIPFMF 130
Db EKNKTQAIYQBYFKYRVCSASTTGELFRFDLDRCPSTEDKVHKEGILLVYKKNIVPYIF 87

Qy 131 RVKRYKVVITSTVYNGIYSD--SITNQHTFYKSIETWETEKMDTIYQCFNSRLNTGCG 188
Db KVRKYKITTYSVRFNGWTRREGVAITNKBLSRAVPKYIEIDMDKTYQCHNCMQIEVNGM 147

Qy 189 LLTVVDRDDNMTVFLQPDVGVTPDVKRYGSQPELYLEPGWFGSYRRRTTVNCELMDMF 248
Db LNSYVDRDGNKTVDLKPVGLTCAITRYISQPKVPADPGWLGTYTRITVACEIVDMF 207

Qy 249 ARSNPPDFVATGDTVMSPFSGEDDHENKMKHPFVSVINNVYKVVQYQNGRVPL 308
Db ARSADPTYFTVTLGDTVEVSPFCVDNCPNATDVLVS--VQIDLNTVVDYGNRATSSQ 265

Qy 309 GKTRIFLDREYTLSEKHLKMSYCPCLTLWKAFYNGIOTESHSGSYHFVANDITASTTS 368
Db HKKIFAFHTLDYSVSEAVNKSASVCSVMFWSQRAIQTEHDLTYHFIANETAGSTV 325

Qy 369 KEDMKFNTVYHCLNEIEKAEIEKKYAKVNSTSKYGDLYKFTDGGLYLVWQPLIQNRL 428
Db KEPLANTSDYNCLMTHINTLEDKIAVNTHTPNGTAYYQTEGGMLVWQPLIAEL 385

Qy 429 LDANKLNNE-----TYRRSRROAESTTDPMMETNGAGGYSSENSITVAQVQYA 481
Db EAMLEATTSPVTPSAPTSSRSKRAIRSIRDV-----SAG-----SENNVLSQIQYA 434

Qy 482 YDNLRIINNILEDLSKAWCEQRAALVWNLKSNPTSVMSMIYNRPVSAKRIGDVIS 541
Db YDKURQSIINNVEELAITWCREQVRQTMVWYEIAKINPTSVMTAIYKPVSRKALGDVIS 494

Qy 542 VSNICVVDQTSVLSHLKSLRLLSASDEKCFGRPPVTFKFMNDSTIYKQGLGVNNEILLTTT 601
Db VTECINVDQSSVSIHKSILK--TENNDICYSRPPVTFKFNSSQLFKQLGARNEILLSSES 552

Qy 602 YLETCQNTYFYFOAKTDMYIKNYEHLKTVPLSSITTLDTFIALNFTLENVDKVIEL 661
Db LVENCHQNAETFTAKNETYHFKNYHVHETLPVNNISTDLTFLALNLTFTIENIDFAVEL 612

Qy 662 YTRDEKLSNVEDLETFRVYNYAQRVSGLRDLDDLVS--TNRNQFVDAFGSLMDLGA 719
Db YSSGERKLANVFDLETFRVYNYAQSISGLRQD--FDNSQRNDRRIIQDFSEILADLGS 671

Qy 720 VQOTVNAVSGVATLFSISVTGFINFKNPFPGGMLMIVIGVLFAIYFLTKTKIYETA 779
Db IGKIVNVASGAPSLFGIVTGILFNKPLPGNFTFLLIGAVILLVILLVRRTNNSQA 731

Qy 780 PIRMIYPEIDKLKREKGSIAPISEBELERIVLAMHIHQONSHMEFTKRDKDSILTR 839
Db PIRMIYPDVEK-----SKSTVTPMEPETIKQILLGMHMQOEAYKKKEQRAARPSIFRQ 786

Qy 840 -AQNMLKRSYSLNKAESV 859
Db AAEFTFLRRSGYKQISTEDKII 807
```

RESULT 2

```
VGLB_EBV
ID VGLB_EBV STANDARD; PRT; 857 AA.
AC P03188;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein GP110 precursor (GP115).
GN BALF4.
OS Epstein-Barr virus (strain B95-8) (Human herpesvirus 4).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Lymphocryptovirus.
OX NCBI_TaxID=10377;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85035713; PubMed=6092825;
RA Bankier A.T., Deininger P.L., Farrell P.J., Barrell B.G.;
RT "Sequence analysis of the 17,166 base-pair EcoRI fragment C of B95-8
  Epstein-Barr virus.";
RL Mol. Biol. Med. 1:21-45 (1983).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=84270667; PubMed=6087149;
RA Baer R., Bankier A.T., Biggin M.D., Deininger P.L., Farrell P.J.,
  Gibson T.J., Hatfull G., Hudson G.S., Satchwell S.C., Seguin C.,
  Tuffnell P.S., Barrell B.G.;
RT "DNA sequence and expression of the B95-8 Epstein-Barr virus genome.";
RL Nature 310:207-211 (1984).
RN [3]
RP IDENTIFICATION OF PROTEIN.
RX MEDLINE=87112946; PubMed=3027378;
RA Gong M., Ooka T., Matsuo T., Kieff E.;
RT "Epstein-Barr virus glycoprotein homologous to herpes simplex virus
  gB.";
RL J. Virol. 61:499-508 (1987).
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
EMBL; V01555; CAA24806.1; -
DR PIR; A03749; QOBEIL.
DR InterPro; IPR000234; Glycoprot B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
KW Glycoprotein; Transmembrane; Signal; Late protein.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 857 GLYCOPROTEIN GP110.
FT DOMAIN 23 688 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 689 704 POTENTIAL.
FT TRANSMEM 710 729 POTENTIAL.
FT TRANSMEM 733 753 POTENTIAL.
FT DOMAIN 754 857 POTENTIAL.
FT CARBOHYD 76 76 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 163 163 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 290 290 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 329 329 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 348 348 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 395 395 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 436 436 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 563 563 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 629 629 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 857 AA; 95639 MW; D9BCE9487D8A1411 CRC64;

Query Match 40.7%; Score 1848.5; DB 1; Length 857;
Best Local Similarity 46.1%; Pred. No. 6.4e-106;
Matches 368; Conservative 148; Mismatches 241; Indels 41; Gaps 12;

Qy 81 STFPYRVCSASGVGVFRFQTDHVCPCD-ASDMVHSEGLLIYKQNIIPFMFRVRYKRV 139
```

```

Db 44 TSPPFRVCLSSGHLDFRSDIQCPSTRENHTEGLMWPKDNIIPYSFKVRSYTKIV 103
Qy 140 TTSTVYNGIYSDITNOHTFYKSEPWETEKMDTIYQCFNSLRNLATNGNLLTYVDRDDIN 199
Db 104 TNILYNGWYASVNRHEEKSVDSYEDQMDTIYQCFNAYVMDKGLTRVYDRDGN 163
Qy 200 MTFVLPQVGVDPDKRYGQSPELYLEPCWFMGYSRRRTTNCVCLMDMFARSNPPDFDV 259
Db 164 ITVNLKPTGGLANGVRRYASQTELYDAPGWLITVTRTTVNCCLITDMMAKSNSPFDFFV 223
Qy 260 TATGTVEMSPWSEDDHENKMKHPFVSINNYKVVDYQNRGTVPGLKTRIPDLREE 319
Db 224 TTGTQVEMSPFYDGK--NKETFERADSFHVRTNYKIVDYDNRGTPQGERAFDLKGT 281
Qy 320 YTLSEKHLKN--MSVCPLTLWKAIFYNGIOTHSYGSHVFVANDITASFSTSKEDMKFEPTT 378
Db 282 YTLSEK--LENRTACPLQHWQTFDSTIATETGKSHFVTDGTSFVNTTNGVIELPDA 339
Qy 379 YHCLNEEIKAETKRYKAKVNSTHSHYKGD--LKYFKTDGGLYLWQPLIQNRLLDKNN--KL 435
Db 340 FKCIIEQVNTKTHKYEAVQDRYTKQEAITYFITSGLLLAWLPLTPRSLATVKNLTTEL 399
Qy 436 NNETHS-----RRSRQAEISTTDPMMETNGAGGEYSSENSI 473
Db 400 TPTTSSPPSPPPAPSAARGSTPAAVLRRRRRRDAGNATTP---VPPTAPGKSLGTLNPP 456
Qy 474 TVAQVQAYDNLRIINLTDLKAWCEQRAALVWNLKINPTSVMSIYNRPVSA 533
Db 457 ATVQIQFAYDSLRRQINRMGLDLARAWCLEQKQNMVRLKINPTTVMSSIIYKAVAA 516
Qy 534 KRIGDIVSNICVVDQTSVSLHSLRLSLASDEKCFRPPVTFKFMNDSTIYKQGLGN 593
Db 517 KRLGDVIVSQCPVQNVQTVTLRKSMR--VPGSETWCYSRPLVSFISFINDTKTYEQQLGTD 575
Qy 594 NEILTTTYTETCQNTYFYQAKTDMYIKNYEHLKTVPLSIITLTDTFIALNPTLLEN 653
Db 576 NEIFLTKMTVEQATQSVYFQSGNEIHVYNDYHFKTIELDGIATLQTFISLNTLSEN 635
Qy 654 VDFKVIELYTRDEKRLSNVEDIETMFREYNYAQRVSGRLKDLLD--LSTNRNQFDARGS 712
Db 636 IDFASELYSRDEQASNVFDELGIFREYNFOAQNIAGLRKDLDNVANGRNQFVDGLGE 695
Qy 713 LMDLLGAVGQTVVNAVSGVATLFSISIVTGFINFKNPFPGMLMIITWIGLVFAIFYLTKK 772
Db 696 LMDLSGSGVQSTNLVSTVGLFSLVSGFISFFKNPFGMLILVLVAGVILVLSLRR 755
Qy 773 TKIYETAPIKMIYPIDKLKEREKSE---TAPISEELERIVLAMHIHQONSHMETKTR 829
Db 756 TRQMSQPVQMLYPGIDELAQOAHSGEGGPGINPISKTELQAIMLA--LHEQNOEQKRAAQ 813
Qy 830 KDPKDSILTRAQNMRLKR 847
Db 814 RAAGPSVASRALQAARDR 831

```

RESULT 3

```

VGLB_RHCM6
ID VGLB_RHCM6 STANDARD; PRT; 854 AA.
AC P89053;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B precursor.
GN GB OR UL55.
OS Rhesus cytomegalovirus (strain 68-1) (RhcMV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=103930;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=97410285; PubMed=9267000;
RA Kropff B., Mach M.;

```

```

RT Identification of the gene coding for rhesus cytomegalovirus
RT glycoprotein B and immunological analysis of the protein.";
RL J. Gen. Virol. 78:1999-2007 (1997).
CC -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; U76749; AAB70024.1; --
DR InterPro: IPR000234; Glycoprot B.
DR Pfam: PF00606; Glycoprotein B; 1.
DR ProDom: PD00693; Glycoprot B; 1.
KW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 23
FT CHAIN 24 854
FT DOMAIN 24 699
FT TRANSMEM 700 720
FT TRANSMEM 724 744
FT DOMAIN 745 854
FT CARBOHYD 40 40
FT CARBOHYD 48 48
FT CARBOHYD 60 60
FT CARBOHYD 183 183
FT CARBOHYD 256 256
FT CARBOHYD 275 275
FT CARBOHYD 314 314
FT CARBOHYD 356 356
FT CARBOHYD 378 378
FT CARBOHYD 382 382
FT CARBOHYD 390 390
FT CARBOHYD 426 426
FT CARBOHYD 426 426
FT CARBOHYD 442 442
FT CARBOHYD 558 558
FT CARBOHYD 595 595
SQ SEQUENCE 854 AA; 97825 MW; F6020DDDE031927A CRC64;

Query Match 24.9%; Score 1134; DB 1; Length 854;
Best Local Similarity 31.5%; Pred. No. 4.6e-62;
Matches 282; Conservative 172; Mismatches 331; Indels 110; Gaps 28;

Qy 20 VALYSLSAETGVTS---PNTATWSTESPLTGHYTHDSSHGGRGNENRDSSEQNKI 76
Db 15 VVYVSIASSSTGTASGAVTPTSPTEITGPII-----ENTLTRHEVFKNM 61
Qy 77 YGSPSTPYRVCSSASGVDFRFTQDHVC---PDASDMVHSEGLIILYKQNIIPFMRV 132
Db 62 ---SKFPYRVCSSMAOQTDLRLFEQINCDSPFKTKEDF--EGIMVYVKRDIRPYTFKV 115
Qy 133 RYKRVKVTSTVYNGIYSDITNOHTFYKSEPWETEKMDTIYQCFNSLRNLATNGNLLTY 192
Db 116 HMYQKILTFQSYSFIRENHLGLFSQEHVAVPMWVHYINKLNCYNSVVRNAG--ATY 173
Qy 193 VD--RDD- INMTVLPQVGVTPDKRYGSOPELYLEPCWFMGYSRRRTTNCVCLMDMFA 249
Db 174 VNYHRDSYVNETMLLVDDDFSNTHSSRFVTVKELMWHKPGSTW-LYTTSCNVCMVTVTTA 232
Qy 250 RSNPPDFFTATGDTVEMSPFWSGEDD--HENKMKHPFVSINNYKVVDYQNRGTVP- 307
Db 233 RSKYPDFFTVTSQGVVDISPFYNGSNKHFGENRDK---FSVRKNYSMTAYYGRDPAE 289
Qy 308 ----LGKTRIFLDREBYTLSEKHLKMSYCPPLTLWKAIFYNGIOTHSYGSHVFVANDIT 362
Db 290 VAHPLVG----FFERPVLMSWDIVEEANNCTEYTFWEQSERTIRSEADTYHTSSMT 345
Qy 363 ASFTTSKEDMKFEFNTTYHCL----NEEIKAEIKKAKVNSTHSHYKGDLYFKTDGGLYL 418
Db 346 ATFLTSKEELNESDPSFQCICKRANEQLQIFNTSY---NETYVQSGVNSMYETTGGLIV 402

```

QY 419 VWOPLIQRLLDAKN---KLNNETYSRRSROAESTTDPMMEMTNGACGEYSSE--NSI 473
 Db 403 FWLPVKEKSIEMEEELAVANNNTSSRRKRSTDSASD-----SNKTSSEVLKSI 452
 QY 474 TVAQOVAYDNLRIRINNILEDLSKAWCREQRAALVWNLKSKINPTSVMSMYYNRPVSA 533
 Db 453 VYAOQLTYTTLRINYINRALRQIAEAWCKQKRTLEVPFKLSKINPMSAMLSAIYDKPIAA 512
 QY 534 KRIGDVISVNCIVVDQTSVSLHKSLLSASDEKCFSPRPVTFKFWNDSTIYKQQLGVN 593
 Db 513 RFVGDVISLAKCDEVDSQVVKLRDMR--TKESGVCYSRPPVLYTFKNSSHVQVQGLGEY 570
 QY 594 NEILLTYYLETQENTYFYFOAKTDMYIYKNVHEHLKTVPLSSITLTDPIALNFTLEN 653
 Db 571 NEILLGHRHTEACEYPSLKIKIYIAGNSSYEVVDVLYKRMIFLDSISTVDTMISLDIDPLEN 630
 QY 654 VDFKVIYLRDEKRLSNVFDIETMPREYNVYAQRVSLGRKDLDLSTNRNQFVDAFGSL 713
 Db 631 TDFKALELYSEDELRSNVFDELDIMREFNTYKQRMITHVEGKVPDKVPG---YLRGLDDM 687
 QY 714 MDDLGAQGVTVNAVSGVATLFSISVTGTFNFKNPFEGG---MLMIIVIGVLPFAIFYLT 770
 Db 688 MSGLSAGKALGVAIGAVGAVASFEVGVGFKNPFSGFTVILFLAVLGVLYIYMRQ 747
 QY 771 KTKIYETAPIKMIYP-----EIDKLERECKSEIAPISSEELERIV 812
 Db 748 KRA--YE-KPHEHFPYVVPPTVKEAPPYSQSYENIKEK-AASATKFSLEEAQOML 803
 QY 813 LAMHIHQONGHMETKTRKDPKD-----SILTRAQNMRLKRSYGYNLKNAESV 859
 Db 804 LA--LQKLDQEKRAADDEDFASNGOSAGFLDLRLN--RRGGYQKIQNEIEV 854

RESULT 4
 ID_VGLB_HCMVT STANDARD; PRT; 907 AA.
 AC P13201;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glycoprotein B precursor [Contains: Glycoprotein GP55].
 GN GB OR UL55.
 OS Human cytomegalovirus (strain Towne).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae; Cytomegalovirus.
 OX NCBI_TaxID=10363;
 RN [1]
 RP MEDLINE=89045645; PubMed=2460994;
 RX Spaete R.R., Thayer R.M., Probert W.S., Masiarz F.R.,
 RA Chamberlain S.H., Rasmussen L., Merigan T.C., Pacht C.;
 RT "Human cytomegalovirus strain Towne glycoprotein B is processed by
 RT proteolytic cleavage";
 RL Virology 167:207-225(1988).
 CC -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M22343; AAA45920.1; -;
 DR PIR; A31288; VGBETE.
 DR InterPro; IPR00234; Glycoprot_B.
 DR Pfam; PF00606; Glycoprotein_B_1.
 DR ProDom; PD000693; Glycoprot_B; 1.
 KW Glycoprotein; Transmembrane; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 907 GLYCOPROTEIN B.

FT CHAIN 461 907 GLYCOPROTEIN GP55.
 FT DOMAIN 25 707 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 708 723 POTENTIAL.
 FT TRANSMEM 729 748 POTENTIAL.
 FT TRANSMEM 752 772 POTENTIAL.
 FT DOMAIN 773 907 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 68 68 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 73 73 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 85 85 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 208 208 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 281 281 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 286 286 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 302 302 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 341 341 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 383 383 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 405 405 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 409 409 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 417 417 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 447 447 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 452 452 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 456 456 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 466 466 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 555 555 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 586 586 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 907 AA; 101953 MW; E6F07B7742D359A2 CRC64;
 Query Match 24.8%; Score 1128; DB 1; Length 907;
 Best Local Similarity 30.9%; Pred. No. 1.2e-61;
 Matches 286; Conservative 184; Mismatches 353; Indels 102; Gaps 25;
 QY 14 LAMVYOVALYSIAETG--VTSPNTATWSTESPLTGHVGHDS--HGERGNNENRDS 69
 Db 5 IWCL--VVCNLCIVCLGAASVSSSTRGTSATHS-----HSSHTTAAHRSRSGVSQRVT 58
 QY 70 EEQ-----NKNIYG-----SPSTFPYVCSASGVGVDFRQTDHVC---PDA 108
 Db 59 SSQTGVSHGVNETIYNTTLKGVGVGVNTTKYPRVCSMAQGTDLIRFERNIVCTSMKPIN 118
 QY 109 SDMVHSEGILLIYQNIIFPMFRVRYKRVKVTSTVYNGIYSDISITQHTFYKSIQWET 168
 Db 119 EDL--DEGIMVYKRNIVAHTEFKRVYQKLTFRRSYAVIHTTYLLGNSNTEYVAPPWMEI 176
 QY 169 EKMDTIYQCFNSLRNLNTGNLLTYVDRDDI--NMTVFLQPDVGVTPDVKVGSPQELYLEP 227
 Db 177 HHINSHSQCYSSYSRVVIAGTVFVAYHRDSEYENKTMQMPDDYSNTHSTRVTVKQDQWHSR 236
 QY 228 GMFWGYSRRRTTVNCELMDMFARSNPPFFVYATGDTVEMSPFSGEDDHENKMEKPW 287
 Db 237 GSTW--LYRETCNLNCWVTITTARSKYPHYHFFATSTGDVVDISPFYNGTNRNASYFGENAD 295
 QY 288 FVSVINNYKVVDYQNRGTVPGLKTRI--FLDREYTLSE--KILKMSYCPPLTLWKAFYN 344
 Db 296 KFFIFPNYTVISDFGRPNSALETHRLVAFLERADSVISMDIQDEKNVT--COLTFWEASER 354
 QY 345 GIQTEHSGYVHFVANDITASFSTTSKEDMKEFNTTYHCLNEEKABIEKY--AKVNSTHSHK 403
 Db 355 TIRSEADSYHSSAKMTATFLSKQEVNMSGALDCVRDEALNKQLQQLFNTSYNQTYEK 414
 QY 404 YGDLKYFTDGGLYLVWQPLIQNLRLDAKNLNNETYSRRSRQAESETTDPMMEMTGNGA 463
 Db 415 YGNVSVEFTTGLVWFVQGIKQKSLVELERLANRSSNLNLTNRKTRST-----DGNNA 467
 QY 464 G--GEYSSSENSITVAQOVAYVAYNLRIRINNILEDLSKAWCREQRAALVWNLKSKINPTS 521
 Db 468 THLSNNMESVHNLVYAOQLQFTYDTRLGYINRALQIAEAWCVDQQRRTLEVFKLSKINPSA 527
 QY 522 VMSMIYNRPVSAKRIGDVISVNCIVVDQTSVSLHKSLLSASDEKCFSPRPVTFKFN 581
 Db 528 ILSAIYNKPIARFMDGVGLASCVTINTQISVKVLRDMN--VKESPGRCSRPPVINFAN 586
 QY 582 DSTIYKQGLGVNNEILLTYYLETQENTYFYFOAKTDMYIYKNVHEHLKTVPLSSITLTD 641
 Db 587 SSVYQYQGLGEDNEILLGNHRTTEECQLPSLKIFLAGNSAYEYVDVLFKRMIDLSSISTVD 646

QY 642 TFIALNFTLENDVFKVIELYTRDEKRLSNVFDIETMREYNVAQVSGRLKDLIDLT 701
 Db 647 SMIALDIDPLENTDFRVLIELSQKRLSSNVFDIEEINRENSYKQRYVEDKVD-- 703
 QY 702 NRNFVDAGSLMDLDGAVGTVNAVSGVATLFSSIVTGFNFINKPFGQMLMIIVIG 761
 Db 704 PLPPYLKGLDDLSGLGAAGKAVGAICAVGAGVAVVEGVATFLKNPFGAFTILVAIA 763
 QY 762 VLFALYFLTKTKIYETAPIKWIYPEI-----DKLXE-----R 794
 Db 764 VIIIIYLIYTORRLCMQPLQNLFPYLVASDGTTVTSNTKDTSLQAPPSYEESVNSGR 823
 QY 795 EG-----KSEIAPISEEELERIVLAM-----HIHQNS-----HMETKTR-KDPKDS 835
 Db 824 KGPQPPSSDASTAPPYNEAQYQMLLALVLDLDAEQRAQQNGTSLDQGTGTQDKGQKPN 883
 QY 836 ILTRAQNLMLRKRGVSNLKNAESVE 860
 Db 884 LLRLRH---RKNGYRHLKDSDEE 905

RESULT 5
 VGLB_HCMVA STANDARD; PRT; 906 AA.
 AC P06473;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glycoprotein B precursor.
 GN GB OR UL55.
 OS Human cytomegalovirus (strain AD169).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae; Cytomegalovirus.
 OX NCBI_TaxID=10360;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87080291; PubMed=3024973;
 RA Cranage M.P., Kouzarides T., Bankier A.T., Satchwell S., Weston K.,
 RA Tomlinson P., Barrell B.G., Hart H., Bell S.E., Minson A.C.,
 RA Smith G.L.;
 RT "Identification of the human cytomegalovirus glycoprotein B gene and
 RT induction of neutralizing antibodies via its expression in
 RT recombinant vaccinia virus.";
 RL EMBO J. 5:3057-3063(1986).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87151139; PubMed=3029980;
 RA Kouzarides T., Bankier A.T., Satchwell S.C., Weston K.,
 RA Tomlinson P., Barrell B.G.;
 RT "Large-scale rearrangement of homologous regions in the genomes of
 RT HCMV and EBV.";
 RL Virology 157:397-413(1987).
 RN [3]
 RP SEQUENCE OF 735-906 FROM N.A.
 RX MEDLINE=87061230; PubMed=3023690;
 RA Kouzarides T., Bankier A.T., Satchwell S.C., Weston K., Tomlinson P.,
 RA Barrell B.G.;
 RT "Sequence and transcription analysis of the human cytomegalovirus DNA
 RT polymerase gene.";
 RL J. Virol. 61:125-133(1987).
 RN [4]
 RP COMPLETE GENOME.
 RX MEDLINE=90269039; PubMed=2161319;
 RA Chee M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,
 RA Horsnell T., Hutchison C.A. III, Kouzarides T., Martignetti J.A.,
 RA Preddie E., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.;
 RT "Analysis of the protein-coding content of the sequence of human
 RT cytomegalovirus strain AD169.";
 RL Curr. Top. Microbiol. Immunol. 154:125-169(1990).
 CC -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL; X17403; CAA35414.1; -;
 EMBL; X04606; CAA28274.1; -;
 EMBL; M17209; AAA46009.1; -;
 EMBL; M14709; AAA45987.1; -;
 PIR; A25365; VGBEC1.
 DR InterPro; IPR000234; Glycoprot B.
 DR Pfam; PF00606; Glycoprotein B; 1.
 DR ProDom; PD000693; Glycoprot_B; 1.
 KW Glycoprotein; Transmembrane; Signal.
 FT SIGNAL 1 24 POTENTIAL.
 FT CHAIN 25 906 GLYCOPROTEIN B.
 FT DOMAIN 25 706 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 707 722 POTENTIAL.
 FT TRANSMEM 728 747 POTENTIAL.
 FT TRANSMEM 751 771 POTENTIAL.
 FT DOMAIN 772 906 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 37 37 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 68 68 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 73 73 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 85 85 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 208 208 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 281 281 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 286 286 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 302 302 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 341 341 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 383 383 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 405 405 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 409 409 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 417 417 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 447 447 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 452 452 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 464 464 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 465 465 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 554 554 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 585 585 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 906 AA; 102003 MW; D4C7A6A3C7083FEE CRC64;

Query Match 24.7%; Score 1121.5; DB 1; Length 906;
 Best Local Similarity 30.4%; Pred. No. 2.9e-61;
 Matches 282; Conservative 186; Mismatches 350; Indels 111; Gaps 25;
 QY 14 LWLYQVALYSLSTAETGVTSPNNTATWSTESPLTGHYTHDSHSGRGNN-ENRD--- 68
 Db 5 IWCL--VVCNLCIVLCGAAVSSSSSTSHATSS-----THNGSHTSRTTSAQTRSVSQ 55
 QY 69 -----SEEQNKNIYG-----SPSTFPYRVCSAGVGDFVRFQTDHVC--- 105
 Db 56 HVTSSSEAVSHRANETIYNTLTLYGDVVGVNTTKYPRVCSMAOQGTDLIRFERNICTSMK 115
 QY 106 PDASDMVHSGILLIYKQNIIPMFRVRKRVKVVTTSTVYNGIYSDITNHTYKSIETP 165
 Db 116 PINEDL--DEGIWVYKRNIVAHFTFKRVYQKVLTFRRSYAYIYTTLLGNSNTYVAPP 173
 QY 166 WETEKMDTIYQCFNSLRNLNTGGNLLTYVDRDDI-NMTVFLOPVDGVTDPVKRYGQSPELY 224
 Db 174 WEIHHINKFAQCYSSRSRVIGTVFVAYHRDSYENKTMQLIPDDYSNTHSTRYVTVKQW 233
 QY 225 LEFGWFGSYRRRTTNCVCELMDFARSNPPDFEFVTATGDTVEMSPFWSGDEDDHKNKHE 284
 Db 234 HSRGSTW-LYRETCNLCMLTITARSKYPHYFATSTGDDVVISPYNGTNRNASYFGE 292
 QY 285 KPMFVSVINNYKVY-DYQNRGTVP-LGKTRIFLDREYTLISWE-KHLKNMSYCPDLTWKA 341
 Db 293 NADKFFIFPNYTVISDFGRNPAAPETHRLVAFERADSVISWDIQDEKNVT-CQLTFWEA 351
 QY 342 FYNGIQTEHSGSYHFVANDITASFTTSKEDMKENTTYHCLNEIEIKAEIEKY-AKNYST 400

```

Db 352 SERTIREADSDFHSFAKTAFLSKQEVNWSDSALDCVRDEANKLQOIFWTSYNOT 411
Qy 401 HSKYGLDKYFKTDGGLYLVWQPLQNLDAKKNLNNETYS---RRSRQAESTTDPME 457
Db 412 YEKYGNVSFETSGGLVFWQGIQKSLVELERANRSSLNTHRTSRSTDDNTHLSS 471
Qy 458 MTGNAGGEVSSSENSITVAQOVAYDNLRIINNILEDLSKAWCREQHRALVWNELSKI 517
Db 472 M-----ESVHNLVYAQLQFTYDLRGVYNRALAQIAEAMCVQORRTLEVFKELSKI 522
Qy 518 NPTSVMSMIYNRPVSARKIGDVISVNCIIVDQTSVSLHKSRLLSASDEKFSRPPVTF 577
Db 523 NPAILSAIYNKPIARFMDGVLGLASCVTINQTSVKVLDMN-VKESPORCVSRPVIF 581
Qy 578 KFMNDSTIYKQGLGVNNEILLTTTYLETQCENTEYFYQAKTDMYIYKNYEHKLTVPULSI 637
Db 582 NFANSSVYVQGLGEDNEILLGNHRTTECOLPSLKIFIAGNSAYEYVDYLFKRMIDLSSI 641
Qy 638 TTLDTFTALNFTLLENVDFKVIELYTRDEKRLSNVFDIETMFREYNYAQRVSLGRKDLL 697
Db 642 STVDSMIALDIDPLENTDFRVELYSQKELRSSNVFDLEEIMREFNSYKQRYVEDKV 701
Qy 698 DLSTNRQVDAFGSLMDDLCAVGQTVVNAVSGVATLFSISIVTGFINFINKNPFGLMII 757
Db 702 D---PLPPYLKGLDLSGLGAAGKAVGAIVGAGVAVSVGEVATFLKNPFGAFTIL 758
Qy 758 VIGVLFAIYFLTKTKIYETAPIKMIYPEI-----DKLKE----- 793
Db 759 VAIATVITLYITRRLCTQPLQNLFFVLVSADGTTVTSGTKDTSLOAPPSYBESVY 818
Qy 794 ---REG-----KSEIAPISEBELERIVLAM-----HIHQONS-----HMETKTP-KD 831
Db 819 NSCRKGCPSPSSDASTAAPTYNEQAYQMLLALRLDAEQRAQOQNGTDSLGGTGTQDKG 878
Qy 832 PKDSILTRAQNLKRSYSLKNAESVE 860
Db 879 QKPNLDRLRH---RKNGYRHLKDSDEE 904

RESULT 6
VGLB_HSV7J STANDARD; PRT; 822 AA.
AC P2352;
AT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B precursor.
GN GB OR U39.
OS Human herpesvirus (type 7 / strain JI) (HHV7).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OX Alphaherpesvirinae; Simplexvirus.
RN NCBI_TaxID=57278;
[1]
RP SEQUENCE FROM N.A.
RA Nicholas J.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. Usage by and for commercial
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; U43400; AAC54701.1; -.
DR PIR; T41941; T41941.
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
DR Glycoprotein; Transmembrane; Signal.
KW

```

```

FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 822 GLYCOPROTEIN B.
FT DOMAIN 23 649 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 650 670 POTENTIAL.
FT DOMAIN 671 683 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 684 704 POTENTIAL.
FT DOMAIN 705 822 EXTRACELLULAR (POTENTIAL).
FT CARBOHYD 152 152 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 244 244 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 324 324 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 338 338 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 352 352 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 358 358 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 387 387 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 482 482 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 517 517 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 567 567 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 822 AA; 93147 MW; F2B410BD359C6498 CRC64;

Query Match 23.9%; Score 1084.5; DB 1; Length 822;
Best Local Similarity 33.6%; Pred. No. 4.8e-59;
Matches 272; Conservative 145; Mismatches 316; Indels 77; Gaps 24;

Qy 78 GSPSTFPYRVCASGVGVDPFQTDHVCBD-ASDMVHSEGLLIYKQNIIPFMRVKYR 136
Db 28 GHQHLPFRICSIATGTDLVRFDFREVCASYSNIIKTTEGILIIYKTKIEAHTFSVRTFK 87
Qy 137 KVVTTSTVYNGI-----YSDSITNQHTFYKSIETPETERKMDTIYQCFNSURLNTG--GNLL 190
Db 88 KELTFQTYRDYDVGTVYFLDRTVTI--TLPMPIE--EVHMVNTTEARCLSSIVKSEEEYV 143
Qy 191 TYVDRODIINMTVFLQVDGVTDPVKRYGQPELYLEPGFWGYSRRRTTVNCELMDMFAR 250
Db 144 AVHKDEYVNTLDLIPNFKSDTVRYITTKPEFLRNGPLW-FYSTSTINCIVTTCIAK 202
Qy 251 SNPPDFPFTATGDTVMSPFPWSGEDDHE-NKMHEKPFWFSVNNY---KVVDYQNRGTV 306
Db 203 TKYPDFPFFALSTGETVEGSPFYNGINSKTFNEPEKILF---RNNTMLKTFDDGSGKNF 259
Qy 307 PLGKTRIFLDREBYTSLWEKHLKNMSYCPLTLWKAIFYNGIQTEHSGSYHFVANDITASFT 366
Db 260 VTLTKMAFLKGNITFISWEQVNEESSICLLKHWMITPHALRAENANSFPIAQLTASFV 319
Qy 367 TSKEDEMKENPTYHCLNEIKABIEKKY-AKVNSTHSKYGDLKYFKTDGGLYLVWQPLIQ 425
Db 320 TGSNTYTLSDSKYNCINSNYTSLIDEIYQTVNNSHDKNGSYEIFKTEGDLIIWQPLIQ 379
Qy 426 NRLLDAKKNLNNETYSRRSRROAESTTDPWMENTGNGAGGEYSSENSITVAQOVAYDNL 485
Db 380 RKLTVLEN-FSNASRRKRR-----ELETNKDIVVQIQLYDTL 418
Qy 486 RIRINNILEDLSKAWCREQHRALVWNELSKINPTSVMSMIYNRPVSARKIGDVISVNC 545
Db 419 KDYINTALGKLAEAWCLNQRTITVLHELKSPSGIISAVYKGPMSAKLIGDVLAVSKC 478
Qy 546 IYVDQTSVSLHKSRLLSASDE---KCFSRPPVTFKPMNDS-TIYKQGLGVNNEILLTTT 601
Db 479 IEVNOTSVQLHKSMLRTKDSYDALRCYSRPLLTYSFANSKSKETIYLGQLGDLNEILLGNH 538
Qy 602 YLETC-QENTEYFYQAKTDMYIYKNYEHKLTVPULSIITLDTFIALNFTLLENVDFKVI 660
Db 539 RTEECQSNKTFLSKGF-AHIFKQTYVNSSLUITEIALDAFVDNLIDPLENADFTLLE 597
Qy 661 LYTRDEKRLSNVFDIETMFREYNYAQRVSGRLKDLLSTNRNQVDAFGSLMDDLCAVG 720
Db 598 LYTKDELKANVDFLETILREYNSYK---SALHHIETKIATVTPTVIGGIDTFPFKGLGAL 654
Qy 721 GQTVNAVSGV-ATLFSSIVTGFINFINKPFGGLMIIIVIGVLFAIYFLTKTKIYETA 779
Db 655 G-LGLGAVLGVTAGALGDVNVGVFSFLKNPFGGALTILLTGLVIGLVFLRHKRLAQT 713
Qy 780 PIKMIYPEIDKLER--EGKSEIAPISEBELERIVLAMHIHQONSHEMETKTRKD--PKDS 835

```

Query Match 23.2%; Score 1055.5; DB 1; Length 830;
Best Local Similarity 32.4%; Pred. No. 2.9e-57;
Matches 245; Conservative 160; Mismatches 292; Indels 59; Gaps 15;

RX MEDLINE=93033157; PubMed=13293336;

CC -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; M97928; AAA43847.1; -;
 DR PIR; A44047;
 DR InterPro; IPR000234; Glycoprot B.
 DR Pfam; PF00606; Glycoprotein_B; 1.
 DR ProDom; PD000693; Glycoprot_B; 1.
 KW Glycoprotein; Transmembrane; Signal.
 FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 24 830 GLYCOPROTEIN B.
 FT DOMAIN 24 651 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 652 674 POTENTIAL.
 FT DOMAIN 675 688 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 689 709 POTENTIAL.
 FT DOMAIN 710 830 EXTRACELLULAR (POTENTIAL).
 FT CARBOHYD 155 155 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 228 228 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 247 247 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 286 286 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 329 329 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 355 355 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 361 361 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 486 486 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 830 AA; 93266 MW; 44B5B8FC6276DC68 CRC64;
 Query Match 23.1%; Score 1050.5; DB 1; Length 830;
 Best Local Similarity 32.3%; Pred. No. 5.9e-57;
 Matches 244; Conservative 160; Mismatches 293; Indels 59; Gaps 15;
 QY 83 FPYRVCASGVGVFRFQTDHVC-PDASDMVHSEGLLIYKONIIPMFVRKRVKVVTT 141
 DB 36 YPFICISIAKGTDLMLREDRDISCPYSKNAKMSGFIIYKTNLETTFPPRYKKELT 95
 QY 142 STVNGIYSISITNQHTFYKSIPEWTEKMDTYQCFNSRLN--TCGNLLTVVDRDIN 199
 DB 96 QSSYRDGVGVYFLDRTVMGLAMPVYEANLVNSHAQCYSAVMKRPDCTVFSAFHEDNKN 155
 QY 200 MTVPLOQVDGTPDVKSYGQPELYLEPCFWGSYRRRTTVNCELMDMPARSPDPFFV 259
 DB 156 NTLNLPFLNFKSITNKRFTTKEPYFARGPLW-LYSTSTSLNCIVTEATAKAKYPFSYFA 214
 QY 260 TATGDTVMSPFWSGED-DHENKMKHPWFVSVNNYKVV-DYQN--RGTVPLGKTRIFL 315
 DB 215 LTGGEIVGSPFFNGSKGFAEPLEK---LTILENTMIEDLMNGMNGATTIVRKIAFL 271
 QY 316 DREBYTLSEKHLKNMSYCLTLWKAFYNGIQTEHSGSYHFVANDITASFTTSKEDMKEF 375
 DB 272 EKADTLFSWEIKENSVCMKHWTTVTHGLRAETDETHYFISKELTAAPVAKPESLNT 331
 QY 376 NTHYCHLNEELKAEIKY-AKVNSTHYSKYGLKYPFTDGLGLYLWQPLQNLDAKNK 434
 DB 332 DPQOTCICKDFEKLINVEVMSYNDYNTSMNGSYQIFKTTGDLILWQPLVQKSLMFL--- 388
 QY 435 LNNETYSRRRROAESTTDPMMETMGAGGEYSSENSITVAOVAYDNLRIINNILE 494
 DB 389 ---EQSEKIRRRDDVV-----DVKSRHDILYVQLYLYDYLKDYINDALG 431
 QY 495 DLSKAWCREQRAALVNNELSKINPTSMVMIMYNRPVSAKRIGDVISVNCIVVDQTSVS 554
 DB 432 NLAESWCLDQKRTITMLHELKSKISPSISIVSEVYGRPISAQLHGDVLAISKIEVNQSSVQ 491
 QY 555 LKSLRLLSA-----SDEKCSRPPVPTKFNWDS-TIYKGLGVNNEILLTTTILETCQEN 609
 DB 492 LKSMRVVDKAGVRSTCMYCNRLPLVTFSEVNSTPEVVPGQLGLDNEILLGDHRTSECEIP 551

QY 610 TEYFQAKTDMYIKNYEHLKTVPLSSITLDTFIALNFTLLENVDFKVIELYTRDEKRL 669
 DB 552 SKTIFLSGNHAYVTDYTHNSTPIEDIEVLDFAFIRKLDPLENADFKVLDLSPDELSR 611
 QY 670 SNVFDIETMPREYNYAQRVSGLRKDLDLSTNRNQFVDAFGSLMDDDLGAVGTQVNVAS 729
 DB 612 ANVFDLENILREYNSYK--SALYTIETAKIATNTPSYVNGINSFLOGLGAIGTGLGSGVIS 668
 QY 730 GVATLFSSIVTGFINKNPFPGGMLMIVVIGLFAIYFLTKTKIYETAPIKWIYP--- 786
 DB 669 VTAGALGDIVGGVSVFLKNPFGGGLMLILAIVVVVIIIVVVRQRHVLKSPIDMFFPYAT 728
 QY 787 -----EIDKLKEREKSEIAPISIEE 807
 DB 729 NPVTTVSSVTGTTVWKTPTSVKVDGGTSVA-VSEKE 763
 RESULT 9
 VGLB HSV6Z STANDARD; PRT; 830 AA.
 AC P36320;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 GN Glycoprotein B precursor.
 GN GB OR U39.
 OS Human herpesvirus (type 6 / strain Z29) (HHV6).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae; Roseolovirus.
 OX NCBI_TaxID=36351;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93033157; PubMed=1329336;
 RA Chou S., Marousek G.I.;
 RT "Homology of the envelope glycoprotein B of human herpesvirus-6 and
 RT cytomegalovirus.";
 RL Virology 191:523-528(1992).
 CC -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; M97927; AAA43846.1; -;
 DR PIR; B44047; B44047.
 DR InterPro; IPR000234; Glycoprot B.
 DR Pfam; PF00606; Glycoprotein_B; 1.
 DR ProDom; PD000693; Glycoprot_B; 1.
 KW Glycoprotein; Transmembrane; Signal.
 FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 24 830 GLYCOPROTEIN B.
 FT DOMAIN 24 651 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 652 674 POTENTIAL.
 FT DOMAIN 675 688 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 689 709 POTENTIAL.
 FT DOMAIN 710 830 EXTRACELLULAR (POTENTIAL).
 FT CARBOHYD 155 155 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 247 247 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 286 286 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 329 329 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 361 361 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 486 486 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 830 AA; 93209 MW; 5EAFE7EC8C43349 CRC64;
 Query Match 22.9%; Score 1040.5; DB 1; Length 830;
 Best Local Similarity 32.3%; Pred. No. 2.4e-56;
 Matches 244; Conservative 155; Mismatches 298; Indels 59; Gaps 15;
 QY 83 FPYRVCASGVGVFRFQTDHVC-PDASDMVHSEGLLIYKONIIPMFVRKRVKVVTT 141


```
Db 36 YPFRICSTAKGDLRFDRDISCSYKSNKSEGFFIYKTNIEYTPPVRTYKNELT 95
Qy 142 STVNGIYSDSITNQHFKYSETEKMDIYOCFNSRLN--TGGNLLTYVDRDDIN 199
Db 96 QTSYRDVGVPFLDRTVGLMPLVYEAANLVNSRAQCYSAVAIKRPGDGVFSAYHEDNNK 155
Qy 200 MTFVLQPDGVTDPVKRYGQSPQLYLEPCGFWGYSYRRRTTNCELMDMFARSNPPDFDV 259
Db 156 ETELEFLFNKSVTNKRFTTKPEYFARGPLW-LYSTSTSLNCIVTEATAKAYFSPFA 214
Qy 260 TATGTVEMSPWSEGD-DHENWMEKHPWFSVINNYKV-DYON--RGTVPGLKTRIFL 315
Db 215 LTTGIVEGSPFPDGSNGKHAEPLEK--LTILENTYMIEDLMNGMGATTLVRKIAFL 271
Qy 316 DREYTLGWEKHLKMWSCPLTLWKAIFYNGIOTEHSGSVHFVANDITASFTSKEDMEF 375
Db 272 EKGDTLFSWEIKEENESVCMKHWTTVTHGLRAETDETHYFISKELTAFAVASKESLNT 331
Qy 376 NTTYHCLNEEIKAEIEKRY-AKVNSTHSGYGLKYFKTDGGLYLVQPLIQNRLLDANK 434
Db 332 DPQKTCIKNEFEKIITDVYMSDYNDYSMNGSYQIFKTTGGDLILIQPLVQKSLM---V 387
Qy 435 LNNETYSRRSROASTTDPMMWMTGNGAGGYSSSENSITVAQVQYAYDNIRIRINILE 494
Db 388 LEQGSVNLRRRRDLV-----DVKSRRHDILYVQLQYLDYTLKDYINDALG 431
Qy 495 DLSKAWCEQHOHRAALVWNLKINPTSVMSIMYNRPVSAKRIGDVISVNCIVVDQTSVS 554
Db 432 NLAESWCUDQKRTITMLHELKISPSISVSEYGRPISAOQLHGDVLAISKIEVNSQSVQ 491
Qy 555 LHKSLRLLSA---SDEKCFRPPVTFKFMNDS-TIYKGOLGVNNEILLTITYLETQEN 609
Db 492 LYKSNRVDAKGVRSCTWYRPLVTFSEVNSTPEVILGQLGDLNEILLGHRTEECIP 551
Qy 610 TEYFQAKTDMYIKXNHEHLKTVPLSSITTLDTFIALNFTLENVDFKVIETYRDEKRL 669
Db 552 STKIFLGNHARVYDYTHNSTPTIEDVLDAFIRLKIDPLENADFLLDLYSPDELSR 611
Qy 670 SNVFDIETMFEYNYVYAOVSGLRDXDLDLSTNRNQFVDAFGSLMDDLGAVGQTVNVAVS 729
Db 612 ANVFLENLIRBYNSYK---SALYTEAKIATNTPSYNGINSFLOGLGAIGTGIGSVIS 668
Qy 730 GVATLFSSIVTGFINFIKPNFCGMLMIWIGVLEFAIVFLTKTKIYETAPIKMYP--- 786
Db 669 VTAGALGDIVGVSVFLKNPFGGGLMLAIIVVVIIIVFVRQKHLVSKPIDMMFPYAT 728
Qy 787 -----EIDKLKEREKGEIAPISEEE 807
Db 729 NPVTTVSSVTGTVVVKTPSVKXADGTSVA-VSEKE 763
```

RESULT 10

```
VGLB_MCMVVS
ID VGLB_MCMVVS STANDARD; PRT; 928 AA.
AC P27171;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-OCT-1993 (Rel. 27, Last annotation update)
DE Glycoprotein B precursor.
GN GB OR UL55.
OS Murine cytomegalovirus (strain Smith).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Muromegalovirus.
OX NCBI_TaxID=10367;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92292266; PubMed=1318410;
RA Rapp M., Messerle M., Buehler B., Tanneheimer M., Keil G.M.,
RA Koszinowski U.H.;
RT "Identification of the murine cytomegalovirus glycoprotein B gene and
RT its expression by recombinant vaccinia virus.";
RL J. Virol. 66:4399-4406(1992).
```

```
RN [2]
RP SEQUENCE OF 1-280 FROM N.A.
RX MEDLINE=93033129; PubMed=1329325;
RA Messerle M., Keil G.M., Schneider K., Koszinowski U.H.;
RT "Characterization of the murine cytomegalovirus genes encoding the
RT major DNA binding protein and the ICP18.5 homolog.";
RL Virology 191:355-367(1992).
RN [3]
RP SEQUENCE OF 514-928 FROM N.A.
RX MEDLINE=92024072; PubMed=1718083;
RA Elliott R., Clark C., Jaquish D., Spector D.H.;
RT "Transcription analysis and sequence of the putative murine
RT cytomegalovirus DNA polymerase gene.";
RL Virology 185:169-186(1991).
CC -!- FUNCTION: IMPORTANT DETERMINANT OF VIRUS INFECTIVITY.
-----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; M86302; AAA45936.1; -
CC EMBL; X67021; CAA47416.1; -
CC EMBL; M73549; AAA45939.1; -
CC PIR; A41992; VGBEMC.
CC InterPro; IPR000234; Glycoprot B.
CC Pfam; PF00606; Glycoprotein B; 1.
CC ProDom; PD000693; Glycoprot B; 1.
CC Glycoprotein; Transmembrane; Signal.
CC SIGNAL 1 36
CC CHAIN 37 928
CC DOMAIN 37 739
CC GLYCOPROTEIN B.
CC EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 740 755
CC POTENTIAL.
CC TRANSMEM 761 780
CC POTENTIAL.
CC TRANSMEM 784 804
CC POTENTIAL.
CC DOMAIN 805 928
CC CYTOPLASMIC (POTENTIAL).
CC CARBOHYD 93 93
CC N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 223 223
CC N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 315 315
CC N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 354 354
CC N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 398 398
CC N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 429 429
CC N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 470 470
CC N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 618 618
CC N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 667 667
CC N-LINKED (GLCNAC. .) (POTENTIAL).
CC CONFLICT 522 522
CC K -> R (IN REF. 3).
CC CONFLICT 527 527
CC D -> E (IN REF. 3).
CC CONFLICT 556 556
CC S -> T (IN REF. 3).
CC CONFLICT 562 562
CC I -> V (IN REF. 3).
CC CONFLICT 583 583
CC K -> Q (IN REF. 3).
CC CONFLICT 587 587
CC S -> K (IN REF. 3).
CC CONFLICT 617 617
CC I -> V (IN REF. 3).
CC CONFLICT 672 672
CC E -> D (IN REF. 3).
CC CONFLICT 675 675
CC D -> Q (IN REF. 3).
CC CONFLICT 806 806
CC R -> RQSAAR (IN REF. 3).
CC CONFLICT 890 890
CC N -> T (IN REF. 3).
CC SEQUENCE 928 AA; 104109 MW; 0EDBF3FE7A43810D CRC64;
Query Match 22.7%; Score 1034; DB 1; Length 928;
Best Local Similarity 31.3%; Pred. No. 7.le-56;
Matches 264; Conservative 155; Mismatches 324; Indels 100; Gaps 25;
```

Qy 8 RGSVLALWLYQVAL-----YLSIAETGVTS-----PNTATWSTESPLT 48

Db 7 RGRSSSWYAMSTALAVTIWCLIACTSEVIAASTPCTTPAKTDTTSETSAETATS 66

Qy 49 GHYTHDSSHGCGNNENRDS-----EQKNKIYGSPTFPYRVCASGVGVDF 97

Db 67 G-----AATGKKEATPTQASKITGTTIVFVNEDMVSVDIDKYPYRVCMAVST-DLV 119

QY 98 RFQDTHVCPDASDMVH-SEGILLIYKQNIIPMPFRVRYKRVKVTSTVYNGIYSDSITNQ 156
 Db 120 RFKGSIDCINHPTKPTQVQEGIMVYKENIVAYFEVITYHK-----DAIFORSVADITINY 175
 QY 157 H---TFYKSIHP-WETEMKPTIYQFNSL-RLNTGGNLLTYVDRDDINMTVFLQPDVGVT 211
 Db 176 FLGTSVTKMAFFIWELDEVNRNRCYSAASRIINGEVVYAHEDSYRNYTMVLVEDDYS 235
 QY 212 PDVKRYGQSELYLEPCWFGWSYRRRTVNCELMDMFARNPDPDPFVTATGDTVMSPP 271
 Db 236 KNSKRYVTTKSRHKGAWTW-RYTESCNMNCVVVVTKARGNTPYEFFVLSSGEVEISPP 294
 QY 272 WSGEDDHNKMKHPFVSVNNYKVDY-----QNRGTVPLGKTRIFLDRREEYTLISWEK 326
 Db 295 YNGENSFPFEDTRNFWR--KNYTKTYFGELAAPKRVPL---MAFLEREDMTIGWEI 349
 QY 327 HLKNMSYCPLTLWKAFYNGIQTEHSGSYHFVANDITASFTTSKEDMKFEFT-----T 378
 Db 350 FPKQNVTCDWKQWTSRAIRTDNTGCHYFVSKGLTATFVASKRKI-DYNTTTEGKNYNT 408
 QY 379 YHCLNEIEKABIEKYYA-KVNSTHYSKYGDLKYFKTDGGLYLVWQPLIQNRLLDKNNKLN 437
 Db 409 FRCVYDFEVEVNRVFEDEYNETHVKDGELEMYRTTGGTLVLWQG-----LKAKSLLNL 462
 QY 438 ETSYRRRRQAEESTDPM-MEMTNG-----AGGEYSSEN-----SITVAQOVYAVDNL 486
 Db 463 EKFAALNNVSATASPPVTTAATENGTVTSRRKRSPDNDDVVVDTSYLAQLOFTYDVLK 522
 QY 487 IRINILEDLSKAWCREQHRRAALVNNELSKINPTSMVMIYRNPVSAKRIGDVISVNCI 546
 Db 523 DYINDALRNIMDAWCRQKRTAEMLKELSKINPSNLSAIYERPVTKLAGDVIANSECV 582
 QY 547 VVDTSVLSKLSLRLLSASDEK---CFSRPPTVFKFNDSTIYKQGLGVNNEILLTTTYL 603
 Db 583 KVDQSSVVKLVKDMRIF--QDGKVVCYSRPLVWPFQFINSTKLESGQJGENNEIMLGTFT 640
 QY 604 ETCQENTEYFQAQTDVMIYKVEHLKTVPLSSITLDTFLALNFTLLENVDFKVIETY 663
 Db 641 ENCDTNRKIFVGVTVGYEYRDRFRNVTSLEHIDLVDTLIGLDIEPLENTDFKVLLEY 700
 QY 664 RDEKRLSNVFDIETMFREYNYYAQRVSLGRKDLDLSTNRNQFVDYAFGLSMDLGAVGQT 723
 Db 701 KGLRASNVFSLDEIMREYNSQKHIRTSAKVDN---NTPSYLLGLDFTMQGLGVAGK 757
 QY 724 VVNAVSGVATLFSISIVTGFINFKNPGGMLIIVIGVLPAIYFLTKTKIYETAPIKM 783
 Db 758 IGVAIGAVGGAVSSVNVAVTGFLTNPFPGFTILLVIGVLAVVLI-----TRPVEY 810
 QY 784 IYP 786
 Db 811 FFP 813
 RESULT 11
 VGLB_HSVT2
 ID_VGLB_HSVT2 STANDARD; PRT; 944 AA.
 AC Q9WRL5;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glycoprotein B precursor.
 GN GB.
 OS Herpesvirus tupaia (Strain 2) (THV-2).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae.
 OX NCBI_TaxID=132678;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9319892; PubMed=10392721;
 RA Bahr U., Springfield C., Tidona C.A., Darai G.;
 RT "Structural organization of a conserved gene cluster of Tupaia
 RT herpesvirus encoding the DNA polymerase, glycoprotein B, a probable

RT processing and transport protein, and the major DNA binding protein.";
 RL Virus Res. 60:123-136(1999).
 CC - SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AF084543; AAD42935.1; --
 DR InterPro: IPR000234; Glycoprot B.
 DR Pfam: PF00606; Glycoprotein_B; 1.
 DR ProDom: PD000693; Glycoprot_B; 1.
 KW Glycoprotein; Transmembrane; Signal.
 FT SIGNAL 1 ? POTENTIAL.
 FT CHAIN ? 944 GLYCOPROTEIN B.
 FT DOMAIN 2 761 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 762 782 POTENTIAL.
 FT DOMAIN 783 785 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 786 806 POTENTIAL.
 FT DOMAIN 807 944
 FT CARBOHYD 114 114 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 238 238 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 369 369 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 409 409 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 414 414 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 426 426 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 481 481 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 485 485 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 620 620 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 944 AA; 106059 MW; 1B063A4CD1561EE2 CRC64;
 Query Match * 22.2%; Score 1008; DB 1; Length 944;
 Best Local Similarity 31.4%; Pred. No. 2.9e-54;
 Matches 260; Conservative 161; Mismatches 322; Indels 84; Gaps 25;
 QY 15 WLYIQA-----LYSLIAETGVSPPTATWSTESPLTGHYGHDSHGERGNENR 67
 Db 41 WYATLIASLVWPTVSSITLLEATVVSSTGGAGQASG--GGGGAGDSTPSE-SPETSA 97
 QY 68 DSEEQNKNYG-----SPSTFPYRVCSASGVDFRFQTDHVC-PDASDMVHSGILL 119
 Db 98 DITVPRVGTGTEWNSLTSERYPYRICSMOGTDIVRFARTITCAPYDAKSVSTGIML 157
 QY 120 IYQNIIPFMFRVRYKRVKVTSTVYNGIYSDSITNQHTFYKSIIEPWETEKMDTIYQCFN 179
 Db 158 IYKANIVPTFDVFTYQKELFFORSYAYIYTVLLGNSREHVALPLWEVDAAINWVCYS 217
 QY 180 SURLNTGGNLLTYVDRDDI-NMTVFLQPDVGVTDPVKRYGQSELYLEPCWFGWSYRRRT 238
 Db 218 SYVRTIGTEQYVSYHQDSYRNETMMLIPEYOSGNTRRYVTVKDQHYVGSTW-LYKETC 276
 QY 239 TVNCELMDMFARNPDPDPFVTATGDTVMSPPW---SGEDDHN--KMKHEKFWFSVIN 293
 Db 277 SMCIVTQTKAKSKYPDYFALSSGLVSEASPPYDVTNGHTTHENRRKHF-----VRE 329
 QY 294 NYKVDVYQNRGTVPGLGKTRI-----FLDREEYTLISWE-KHLKNMSYCPLTLWKAFYNGIQ 347
 Db 330 QYRMLERFCAVNAVPV---RVVPKMAFLERPDIVLAWIEKPKNVT-CHLALWETVNAIR 385
 QY 348 TEHSGSYHFVANDITASFTTSKEDMKFEFTT-YHCLNEIEKABIEKYYAKVNSTHYSKY-- 404
 Db 386 TEHATSFHFVSRLGTATFTVAKANETLYNNSRYDCIRDSANRTIDRVFRE--BYDGKYL 443
 QY 405 -GDLKYFKTDGGLYLVWQPLIQNRLLDKNNK-----NNETYSRRSRQAEESTTDPWME 457
 Db 444 DGDVPIFTTNGGLTVVWQGLRQ-KAALASALAGIPGANGTTNHSRRRD----- 492
 QY 458 MTGNGAGGYSSENSITVAQOVYAVDNLRIINILEDLSKAWCREQHRRAALVNNELSKI 517

```

Db 493 -TAAIAREHASD--LTYAQOFAYDTIRDVYNQAIGHIAEAWCUEQRRTGEMLHLSKI 549
QY 518 NPTSMVMYINRPSVAKRICDVISVNCIWDQTSVSLHSLRLLSAS-----DEKCF 570
Db 550 NPTSMVMYINRPSVAKRICDVISVNCIWDQTSVSLHSLRLLSAS-----DEKCF 570
QY 571 SPPVTFKPMNDSTYKYGOLGVNNEILLTTTLETQCENTEYFFQAKTDMYTKYNEHLK 630
Db 610 SRPVLFREVNSETQYQOLGEDNEILLGTFRTACQLPSLKIFVAGKVAYEYRDLYKR 669
QY 631 TVPLSSITLDTFIALNFTLLNVPKVELYTRDEKRLSNVPDIETMREYNYAQRVS 690
Db 670 QIDLOSIDVNTMISLKVEPLENTDFQVLELSYRGELKSANYFDEIDIMREYNAKRLR 729
QY 691 GLRKDLIDLSTNRNFVDAFGSLMDLGAQVQ---TVNAVSGVATLFSSIVTGFNFIK 747
Db 730 YITKIVNPIP---PFMRGLDDFMSGLGAAGKGLGLVLGAVGAV---ASVVGGSFPT 783
QY 748 NPTSMVMYINRPSVAKRICDVISVNCIWDQTSVSLHSLRLLSAS-----DEKCF 570
Db 784 NPTSMVMYINRPSVAKRICDVISVNCIWDQTSVSLHSLRLLSAS-----DEKCF 570

RESULT 12
VGLB_GPCM
ID VGLB_GPCM STANDARD; PRT; 901 AA.
AC Q69024;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B precursor.
GB GB OR UL55.
OS Guinea pig cytomegalovirus (strain 22122 / ATCC VR682) (GPCMV).
OC Viruses; daDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Muromegalovirus.
OX NCBI_TaxID=103920;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94279138; PubMed=8009831;
RA Schleiss M.R.;
RT "Cloning and characterization of the guinea pig cytomegalovirus
glycoprotein B gene.";
RL Virology 202:173-185(1994).
CC -! SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L25706; AAA4831.1; -
DR InterPro; IPR000234; Glycoprot B.
DR Pfam; PF00606; Glycoprotein B; 1.
DR ProDom; PD000693; Glycoprot B; 1.
KW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 34 POTENTIAL.
FT CHAIN 35 901 GLYCOPROTEIN B.
FT DOMAIN 35 713 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 714 734 POTENTIAL.
FT TRANSMEM 737 757 POTENTIAL.
FT DOMAIN 758 901 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 60 60 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 66 66 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 270 270 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 328 328 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 372 372 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 398 398 N-LINKED (GLCNAC. . .) (POTENTIAL).

```

```

FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 436 436 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 537 537 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 571 571 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 623 623 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 901 AA; 102239 MW; 43ABCA7D45A197AA CRC64;

Query Match 22.0%; Score 999; DB 1; Length 901;
Best Local Similarity 29.1%; Pred. No. 9.6e-54;
Matches 271; Conservative 159; Mismatches 356; Indels 144; Gaps 25;

QY 24 SLSTIAETGVTSPNTATWSTESPLTGHYGHYTHDSSHGERGNENRDSSEONKNI-----YG 78
Db 24 ALTILYLGVCESSTVTPTTVEDTTVSNNGNH--SDASRNNVTIR-----NLTSASVDFS 74
QY 79 SPSPFPYVCSAGVGVDFRQTDHVCDDASDMVHS--EGILLIYKQNIIPNFRVRKYRK 137
Db 75 QRKLYPYRISCMGMDTLVRFARTIQCVFPNPRVNSEEGIMLIYKRNILPVYFTAYTQK 134
QY 138 VVTSTVYNGIVSDSITNQHTFYKSIPEWTEKMDTIYQCFNSLRNLNTGGN--LLTYVDORD 196
Db 135 ELLFORSYKYVTDYLLGYRSEFVALPMWEFLVNSRCQCYTSHORVIGADRYIAHNDN 194
QY 197 DINMTVFLQPDGVTPDVRYGSOPELYLEPGWFMGYSYRRRTTNCMLDMFARNPPPD 256
Db 195 EVNETMMLMRDDMGNDTTRYITVKEHARTPGSVN--LYKETCSMNCIVTKTKGSKFPYD 253
QY 257 FVVTATGDTVEMSPWSGED-----DHNKMKHEKPMFVSVINNYKV--DYQNRGVTPGK 310
Db 254 MFVLPSGVIVNISPFYNGSNGKTFREQEKPH-----IWSNSILKDFGRAL-----E 302
QY 311 TRI-----FLDREEVYLSWEKHLKMSYCPULLWKAIFYNGIQTEHSGSVHFVANDITASF 365
Db 303 ARIYKMAFYEREDVIGWVNDGNSVTCMILMETVDRAIRTEYENAFHYVARTLTSTF 362
QY 366 TTSC-----EDMKFNTTYHCLNBEIKAEIEKKYAK--VNSTHSGYKGLKYFKTDGG 415
Db 363 VENKYSPPNNLTEDDIK-----CFKNDQAOKIEEVFLRDYDNYTDMGNATYHVTGG 415
QY 416 LYLWQVPLIQRLDQAKNLNETHYR--RSRQRAESTTDDMMETNGAGGSEYSENSI 473
Db 416 LVIWQGLKQKSLKALEIAANESAVSATGNSRRKSLPD---ESTGD-----I 461
QY 474 TVAQOVAYDNLIRINNILEDLSKAWCREOHRALVWNLKINPTSVMSYINRPSVA 533
Db 462 SYAQOFAYDTRYINQALGHIAEAWCLODKRTAEVLHLSKINPILSALFGVPVAA 521
QY 534 KRIGDVISVNCIYVDQTSVSLHSLRLLSASD--EKFSRPPVTFKPMNDSTIYKGLQ 591
Db 532 RVVGDVISLAKIEVNQSTVLIKGMKRFSDGKLEGYCSRPVWFMKNSTEVRLGQLG 581
QY 592 VVNEILLTTTLETQCENTEYFFQAKTDMYTKYNEHLKTVPLSSITLDTFIALNFTLL 651
Db 582 EDNEILLGTHRMETCQTDYRIFVAGDYGVEFQYVFTKINLSEIDIITMIALKTEPL 641
QY 652 ENVDPKVIELYTRDEKRLSNVPDIETMREYNYAQRVSGRLKOLLDLSTNR--NQFYDA 709
Db 642 ENIDPKVLELYSRDELAQANVFDESIMREYNYOKKR-----LDFVVERVINPIPPA 693
QY 710 FGSL---MDDLGAVGQTVNAVSGVATLFSSIVTGFNFIKPNFGGLMLIWIIGVLF 766
Db 694 LKGLDEMNGMGAIGKIGEAAGVAGVGAIGSFGALVTFTVNPFGAFVVFVFCVGCITLV 753
QY 767 YFLTKYKIYETAPIKMIYPEIDK-----LKEREGKSEIAP----- 802
Db 754 ITVYRRORRAMQRPDFFFPYASQITSSVADSSIAVAYPGEGTSGDAPPYGEAPYG 813
QY 803 -----ISEELERLVLAWHIHOQNSHMETKTRDKPKDSILTR- 839
Db 814 YKOLSVADTRVSSSSAGAGAFNEEDAVRMLRA--TKRLDDKKRQETEKSSKDSASKN 871
QY 840 -----AQNMLKRKSGYSLNKAESVEM 861

```

```

Db      872 SETRRRPGIMDRLLRRRGYQKLNTEDDVHV 901

RESULT 13
VGLB_HSVSM STANDARD; PRT; 920 AA.
AC Q0464;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B precursor.
GN U127.
OS Herpesvirus saimiri (type 1 / strain MV-5-4-PSL) (Marmoset
herpesvirus).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae.
OX NCBI_TaxID=10353;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93229440; PubMed=8385913;
RA Eberle R., Black D.;
RT "Sequence analysis of herpes simplex virus gB gene homologs of two
RT platyrrhine monkey alpha-herpesviruses.";
RL Arch. Virol. 129:167-182(1993).
CC -1- SUBUNIT: DIMER, PROBABLY LINKED BY DISULFIDE BONDS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- MISCELLANEOUS: GB IS THE ONLY GLYCOPROTEIN THAT IS KNOWN TO BE
CC REQUIRED FOR VIRAL GROWTH.
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M95786; AAA43841.1; -
DR InterPro; IPR00234; Glycoprot B.
DR Pfam; PF00606; Glycoprotein B; 1.
DR ProDom; PD000693; Glycoprot B; 1.
KW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 28
FT CHAIN 29 920 GLYCOPROTEIN B.
FT DOMAIN 29 739 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 740 755 POTENTIAL.
FT TRANSMEM 761 781 POTENTIAL.
FT TRANSMEM 784 804 POTENTIAL.
FT DOMAIN 805 920 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 119 119 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 152 152 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 409 409 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 683 683 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 733 733 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 920 AA; 101661 MW; 62C4D0315D8C2DB7 CRC64;

Query Match 19.8%; Score 900; DB 1; Length 920;
Best Local Similarity 29.0%; Pred. No. 1.2e-47;
Matches 243; Conservative 154; Mismatches 359; Indels 82; Gaps 21;

Qy 60 ERGNENRDS-----EEQNKVIYSPSTFPYRVCSAGVGDFRQTDHVCPCD-ASDMVHS 114
Db 95 QRANKTRDARAQLRSVRQIRAEANATSMFYVCPPTGATVWQFEEPRPCPDVAAGKNET 154
Qy 115 EGILLIYKQNLIPMFVRVYRKVVTSTVY-----NGIYSDSITNGHTYKSLPEW 166
Db 155 EGIAVIFKRNIAPKYFATWYKEITVTQWGSRYLQLTGLYND-----RAPVPP 205
Qy 167 E--TEKMDTIYQCFNSLRNLGTGNNLTLYVDRDDINMTVFVLPQVDGVTDPVKRYGSOPELY 224

```

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

[illegible]

```
QY 2 AGSLKRGSLALWLYQVALYSLSIAETGVT-----SPPTATWTSTESPLTGHY 51
Db 111 MVHSEGLLLYKONIIFPFRVRYKRVKVTITVYNGIYSDSITNQHTFFKYSIEPWE-TE 169
QY 41 AGGARAALAAALLWATWALLAPAGRPATTPAPPEEAAAPAPPASPSPGPGDDDA 100
Db 52 GTDSSHGGERNNENRDEBQNKNIYGSPPFPYRVCSAGSGVDVFRFOTDHHVCPD-ASD 110
QY 101 ASPDNSTDVRAALRLAQAAGENSRRF-----VCPPPSGATVVRADARPCEPYGLG 151
Db 111 MVHSEGLLLYKONIIFPFRVRYKRVKVTITVYNGIYSDSITNQHTFFKYSIEPWE-TE 169
QY 152 RNYTEGIGVLYKENIAPYTFKVIYKRVIVITWAGSTYA-AITNQYTDVVPVGMGEITD 210
Db 170 KMDTIYOCFNSRLNTCGNLLTVDRDDINMTVFLQPDVGVTPDVKRYGSOPELYLEPGW 229
QY 211 LVDKKWRCLSKAEVLSRGRKVAFDRDDDPWEAPLKPALSAPGVRGWHTTDDVYALG- 269
Db 230 FWSGRRRTTNCGLMDMFARSNPPDFFTVATGDTVMESPFW-SGDDHENKMKHPWF 288
QY 270 SAGLYRTGTSVNCIVEEVEARSVPYDSFALSTGDIITYMSPFYGLREGAHEHTSYSPER 329
Db 289 VSVNNYKVVDYQNRGTVPLGKTRIFLDREYTLNWEKHLKNMYSYCPDLTLWKAFYNGIQ 348
QY 330 FOLEGYKXKDMATGRRLKPEVSRNFRLOHVTAMDVMPKRVKNCVSLAKWREADEMLRD 389
Db 349 EHSQSYHFVANDITASFTTSKEDMKFNFTY-HCLNEEIKAEIEKKY-AKVNSTHSGYGD 406
QY 390 ESRGNFRFTARSLSATFVSDSHTFALQNVPLSDCVIEEAAEAAVERVYREYNGTHVLSGS 449
Db 407 LKYFKTDGGYLVWQPLIQRLLDANKL-----NNETYS-----RRSRQAE 449
QY 450 LETYLRGGFWAFRPMLSNEL--AKLYQLARSNGTLEGLFAAAAPKPGPRARRAAP 507
Db 450 STTDPNMWMTGNG-----AGGEYSSENSITVAQOVAYDNLIRINNILEDLSKACRE 503
QY 508 SA--PGGFGAANGPAGDGDAGGRVTTVSSAEFAALQFYDHIQDHVNTMFSRLATSWCLL 565
Db 504 QHRAALVWNELSKINPTSVMSMIYNRVPSAKRIGDVISVNCIVDQTSVSLHKSRLLS 563
QY 566 QNKERALWAEAAKLNPSAASAAALDRRAAARMGLGDAMAVTYCHELGEGRVFIENSNR- 622
Db 564 ASBEKCFRPPVTFKFNWSTIYKQGLGVNNEILLTLYTLETQENTYFYFQAKTDMYIY 623
QY 623 APGVCYCRPVPVFAFNGSEPEVEGQLGEDNELLPGRELVEPCTANHCRYFRFCADVYV 682
Db 624 KNYEHLKTVPLSSITTLDTIALNFTLLENVDKVIYELTYRDEKRLSNVFDIETMPREYN 683
QY 683 ENYAVRVRPLAELEVISITFDVNLNLTVLEDRFEPLPLEVYTRAEADTGLLDYSEIORNQ 742
Db 684 YYAQRVSGLRKDLIDLSTNRNQFVDFAGSLMDLDLGAVGQTVWNAVSGVATLFSISVTGFI 743
QY 743 LHELRFYDI--DRVVKTDGNMIMRGLANFFOGLGAVGQAVGTIVLGAAGAALSTVSGIA 800
Db 744 NFKNPPGMLMIIVTGVLFALYFLPKTKIVETAPIKMIYEPIDKLKEREKSEIAP 803
QY 801 SFIANPEGALATGILLVAGVAAFLAYRYISRLSRNPMKALYPTITRAALKDDARGATAP 860
Db 804 SESE-----LER---IVLAWHIHQONSHMETKTRDKPDS--ILTR-AQNMLRKSS--G 849
QY 861 EEEEFDAALQAREMIKYMSLVSAVEROEKAKKSKNGGPLLATRLTOLALRRAPPE 920
Db 850 YSNLKNAE 857
QY 921 YQQLPMAD 928
```

RESULT 16

```
VGLB_HSV1K
ID_VGLB_HSV1K STANDARD; PRT; 904 AA.
AC P06437;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B precursor.
```

```
GN OR UL27.
OS Herpes simplex virus (type 1 / strain KOS).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10306;
RN [1]
RX SEQUENCE FROM N.A.
RX MEDLINE=84174058; PubMed=6324454;
RA Bzik D.J., Fox B.A., Deluca N.A., Person S.;
RT "The nucleotide sequence of the glycoprotein gene, gB, of herpes simplex virus type 1.";
RL Virology 133:301-314(1984).
RN [2]
RX REVISIONS. SEQUENCE FROM N.A.
RX MEDLINE=87071654; PubMed=3024391;
RA Bzik D.J., Debroy C., Fox B.A., Pederson N.E., Person S.;
RT "The nucleotide sequence of the gB glycoprotein gene of HSV-2 and comparison with the corresponding gene of HSV-1.";
RL Virology 155:322-333(1986).
RN [3]
RX REVISIONS.
RA Pederson N.E.;
RL Submitted (APR-1987) to the EMBL/GenBank/DBJ databases.
CC -!- SUBUNIT: DIMER, PROBABLY LINKED BY DISULFIDE BONDS.
CC -!- MISCELLANEOUS: THERE ARE SEVEN EXTERNAL GLYCOPROTEINS IN HSV1: GH,
CC GB, GC, GG, GD, GI, AND GE.
CC -!- MISCELLANEOUS: GB IS THE ONLY GLYCOPROTEIN THAT IS KNOWN TO BE
CC REQUIRED FOR VIRAL GROWTH.
CC -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch)
CC -----
DB EMBL: K01760; AAA45774.1; --
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
KW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 30
FT CHAIN 31 904 GLYCOPROTEIN B.
FT DOMAIN 31 730 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 731 746 POTENTIAL.
FT TRANSMEM 752 771 POTENTIAL.
FT TRANSMEM 775 795 POTENTIAL.
FT DOMAIN 796 904 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 398 398 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 430 430 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 489 489 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 674 674 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 904 AA; 100368 MW; B97B7F8DE5FBA299 CRC64;
Query Match 18.7%; Score 849.5; DB 1; Length 904;
Best Local Similarity 27.7%; Pred. No. 1.4e-44;
Matches 245; Conservative 182; Mismatches 358; Indels 101; Gaps 28;
QY 25 LSIAETGVTSPTATWSTESP--LTGHYGHYDSSHGE-RGNENRDSSEQNKNIYGPS 81
Db 59 LGAAPTGDPKPKKKNKPNTPPRPADGNATVAAGHATLREHLRDKAENTDANFY---- 114
QY 82 TFPYRVCSAGSGVDVFRFOTDHHVCPDASD-MVHSEGLLLYKONIIFPFRVRYKRVKVT 140
Db 115 -----VCPPTGATVVQFQPRCPTRPQGVNTEGIAVFKENIAPYKFKATMYKQVT 169
QY 141 TSTV-----YNGIYSDSITNQHTFFKYSIEPWE--TEKMDTIYOCFNSRLNTGNLL 190
Db 170 VSQVWFGRHYSQFMGIFED-----RAPVFEEVIDKINAKGVCRTAKYVRNLET 220
```

```

QY 191 TVVDRDDINMTVFLQPVGV-----TPDVKRYGSOPELYLEPGWFMWGSYRRRTTVC 242
Db 221 TAFHRDDHETDMLKPANAATRTSRGWHHTDLKYNPSRVEAF-----HRYGTTVNC 271
QY 243 ELMDFARSNPPDFEFTVATGDTVMSPFWS-GEEDH-ENKMHKPFVSVINNY-KVVD 299
Db 272 IVEEDARSVPYDFEVLATGDFVMSPPFYGYREGSHTHTTYAADRFKQVDGFGYARDLT 331
QY 300 YONRGTGVLGKTRIFLDREYVTLSEKHLKMSYCPJTLKAFYNGIOTHSQSVHFVAN 359
Db 332 TKARATAP--TTRNLLTTPKFVANDWVPKRSVCTMTKWQVDEMLSEYGGSPRFSDD 389
QY 360 DITASFTT-----SKEDMKFNTTYHCLNBEIEKKA-YNVSTHSHYKGLDKYFK 411
Db 390 AISTFTTNTLTPELSRVDLGD-----CIGKARDAMDRIFFARRYNATHIKVGQPYQ 443
QY 412 TDGGLYVWQPLQNRLLDAKNNLNNYTSRRSRQAESTTDPMMWMTGNAGGYSSEN 471
Db 444 ANGSGFLIAYQPLLSNTLAE---LYVREHLRQSRKPPNPTPPPGASANASVERIKTTS 499
QY 472 SITVAQOVAYDNLIRINNILEDLSKACRQHRALVNMNLSKINPTSVMSIYNRPV 531
Db 500 STEFARLQFTYHIOGHVNDMLGRVIAIACELQNHLELTMNEARKLNPNATASVTVGRV 559
QY 532 SAKRIDGVISNCIIVDPQTSVSLHKLRLSASDEKCFSPRPVTFKFMNDSTIYKGLQ 591
Db 560 SARMLGDVMAVSTCVPAADNVIVQNSMR-ISSRPGACYSRPLVSFRYEDQGLPVGQ 618
QY 592 VNNEILLTTLTLETCQENTYFQAKDMYIYKNYEHKLKTPVPLSSITLDTFIALNFTLL 651
Db 619 ENNELRLTRDAIEPCVTGHRHYFTFGGVYVPEEYAYSHQLSRADITTVSTFDLNI 678
QY 652 ENVDKVELYTRDRKSLNSVDFDIETMREYNVYAAQVSLGRKLDLSTNRNQP--VDA 709
Db 679 EDHEFVPLEVYTRHEIKOSGLLDYTEVQRRNQLDLRFADI-DTVIHADANAAMFAGL 737
QY 710 FGSLMDDLQ-AVGQTVVNAVSGVATLFSISVITGFINFKNPFGGMLIIVIGVLFALYF 768
Db 738 PFEGMGDLGRAVKVMGIVGVW---SAYSGVSSFSNPFALAGVLLVLGLAAFAFF 793
QY 769 LTKTKIYETAPIKMIYPEIDK-LK-----EREGKSEIAPISE-BELERIVUAM 815
Db 794 AFYVYVRLQSNPKALYPLTTKELKNPTNPASGEGEGGDFDEAKLAEREMIRYMALV 853
QY 816 HIHQNSHMETKTRDPKDSILT-RAQNML---RKESGYNLKNAE 857
Db 854 SAMERTEH---KAKKKGTSALLSAKVTDMVRKRRTNTYQVPPNKD 896

RESULT 17
VCLB_HSV1F STANDARD; PRT; 903 AA.
AC P06436;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B precursor.
GN GB OR UL27.
OS Herpes simplex virus (type 1 / strain F).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10304;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85083254; PubMed=2981343;
RA Pellett P.E., Kousoulas K.G., Pereira L., Roizman B.;
RT "Anatomy of the herpes simplex virus 1 strain F glycoprotein B gene:
RT primary sequence and predicted protein structure of the wild type and
RT of monoclonal antibody-resistant mutants.";
RL J. Virol. 53:243-253 (1985).
RN [2]
RP SEQUENCE OF 1-176 FROM N.A.

```

```

RX MEDLINE=88306232; PubMed=2457278;
RA Hammerschmidt W., Conraths F., Mankertz J., Buhk H.-J., Pauli G.,
RA Ludwig H.;
RT "Common epitopes of glycoprotein B map within the major DNA-binding
RT proteins of bovine herpesvirus type 2 (BHV-2) and herpes simplex
RT virus type 1 (HSV-1).";
RL Virology 165:406-418 (1988).
CC -!- SUBUNIT: DIMER, PROBABLY LINKED BY DISULFIDE BONDS.
CC -!- MISCELLANEOUS: THERE ARE SEVEN EXTERNAL GLYCOPROTEINS IN HSV1: GH,
CC GB, GC, GG, GD, GI, AND GE.
CC -!- MISCELLANEOUS: GB IS THE ONLY GLYCOPROTEIN THAT IS KNOWN TO BE
CC REQUIRED FOR VIRAL GROWTH.
CC -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M14164; AAA45776.1; --
CC EMBL; M21633; AAA45788.1; --
CC PIR; A03750; VGBEB1.
CC InterPro; IPR000234; Glycoprot B.
CC Pfam; PF00606; Glycoprotein B_1.
CC ProDom; PD000693; Glycoprot_B_1.
CC Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 29
FT CHAIN 30 903 GLYCOPROTEIN B.
FT DOMAIN 31 729 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 730 745 POTENTIAL.
FT TRANSMEM 751 770 POTENTIAL.
FT TRANSMEM 774 794 POTENTIAL.
FT DOMAIN 795 903 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 86 86 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 140 140 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 397 397 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 429 429 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 488 488 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 673 673 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 903 AA; 100104 MW; 73BDCA7813DB35E8 CRC64;

Query Match 18.6%; Score 845.5; DB 1; Length 903;
Best Local Similarity 27.7%; Pred. No. 2.5e+44;
Matches 245; Conservative 179; Mismatches 358; Indels 101; Gaps 28;

QY 28 AETGVTSPPTATWSTESP--LTGHYGHDSHGE-RGNENRDSSEQNKNYGSPTFP 84
Db 61 APTGDTKPKKKPKPNPPPPAGDNATVAAGHATLREHLRDKAENTDANFY----- 113
QY 85 YRVCSASGVGDVFRFQTDHVCPPDASD-MVHSEGLLIYKQNIIPFMRVRKYRVVTTST 143
Db 114 --VCPPTGATVVGFEQPRCPTREQGNVTEGIAVFKENIAPYKFKATMYKDVTVSQ 171
QY 144 V-----YNGIYSDSITNQHTFYKSIQPEW--TEKMDTIYQCFNSLRNLGNTLYV 193
Db 172 VWFGRHYRQFMGIFED-----RAPVPFEEVIDKINAKGVCRTAKYVRNNLETTAF 222
QY 194 DRDDINMTVFLQPVGV-----TPDVKRYGSOPELYLEPGWFMWGSYRRRTTVCNCLM 245
Db 223 HRDDHETDMLKPANAATRTSRGWHHTDLKYNPSRVEAF-----HRYGTTVNCIVE 273
QY 246 DMFARSNPPDFEFTVATGDTVMSPFWS-GEEDH-ENKMHKPFVSVINNY-KVVDYQN 302
Db 274 EVDARSVPYDFEVLATGDFVMSPPFYGYREGSHTHTTYAADRFKQVDGFGYARDLTKA 333
QY 303 RGTVPCLKTRIFLDREYVTLSEKHLKMSYCPJTLKAFYNGIOTHSQSVHFVANDIT 362
Db 334 RATAP--TTRNLLTTPKFVANDWVPKRSVCTMTKWQVDEMLRSEYGGSPRFSDAIS 391
QY 363 ASFTT-----SKEDMKFNTTYHCLNBEIEKKA-YNVSTHSHYKGLDKYFKTDG 414

```

```
Db 392 TTTTNTLTPLSRDLGD-----CIGKDARDMDRIFARRYNATHIKVGQOYYLANG 445
QY 415 GLYLVQPLQIQLLDAKNNKNNETYSRRRQAESTTDDPMEMTNGAGGEYSSENSIT 474
Db 446 GFLLIAYOPLLSNTLAE-----LYVREHLREQSRKPPNTPPPPCASANASVERIKTTSSIE 501
QY 475 VAQOYAYDNLRIINILLEDLSKAMCREQHRAALVWNLKSKINPTSVMSMIYNRPVSAC 534
Db 502 FARLOQTYNHIQRHVNDMLGRVAIAWCELQNHLETLWNEARKLNPNNAIASATVGRVSAR 561
QY 535 RIGDVTLSVNCIIVDQTSVLSHLKSLRLSDEKSCFSPVPVTFKFWNDSTIYKQGLGVNN 594
Db 562 MLDGWAIVSTCVPAADNVIVQNSMR--ISSRPGACYSRPLVSFRYEDQGLVEGQLGENN 620
QY 595 EILLTTTYLETCTENTYFYQAKTDMYIYKNYHLKTVPLSSITTLDTFIALNFTLLENV 654
Db 621 ELRLTRDAIEPCTVGHRRYFTFGGYYVPEYAYSHQLSRADITTVSTFDLNLITWLEDH 680
QY 655 DFKVIELYTRDEKRLSNVFDIETMFREYNYAORVGLRKLRLDLSTNRQF--VDAFGS 712
Db 681 EFPVLEYVYTHEIKDGLDYTEVQRNQLHDLRFAD1-DTVIHADANAAMFAGLGAFFE 739
QY 713 LMDDLG-AVGQTVNAVSGVATLFSSTVTCFINFKNPGCGMLMIIVIGVLPFAIVFLTK 771
Db 740 GMDGLGRAVKVWGVGVV-----SAVSGVSFMSNPPGALAVGLLVLAGLAAAFAPR 795
QY 772 KTKIYETAPIKMIYPEIDK-LK-----EREKGSBIAPISE-EELERIVLAMHIH 818
Db 796 YVMRLQSNPMKALYPLTTKELKNPTNPDSAGEGEGGDFDEAKLAEMIRYMALVSAM 855
QY 819 QQNSHMETKTRDKPKDSILT-RAQNML---RKSGSYNLKNAE 857
Db 856 ERTEH---KAKKXGTSALLSAKVTDVMVRKRNTNTYQVFNKD 895

RESULT 18
VGLB_HSV2H STANDARD; PRT; 904 AA.
AC P08666; P89450;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B precursor.
GN GB OR UL27.
OS Herpes simplex virus (type 2 / strain HG52).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10315;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87071654; PubMed=3024391;
RA Bzik D.J., Debroy C., Fox B.A., Pederson N.E., Person S.;
RT "The nucleotide sequence of the gB glycoprotein gene of HSV-2 and
RT comparison with the corresponding gene of HSV-1.";
RL Virology 155:322-333(1986).
RN [2]
RP SEQUENCE FROM N.A.
RA Dolan A.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: DIMER, PROBABLY LINKED BY DISULFIDE BONDS.
CC -1- MISCELLANEOUS: THERE ARE SEVEN EXTERNAL GLYCOPROTEINS IN HSV1 AND
CC 2: GH, GB, GC, GD, GI, AND GE.
CC -1- MISCELLANEOUS: GB IS THE ONLY GLYCOPROTEIN THAT IS KNOWN TO BE
CC REQUIRED FOR VIRAL GROWTH.
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/)
CC
```

```
or send an email to license@isb-sib.ch.
CC
CC EMBL; M14923; AAA66440.1; -.
DR EMBL; Z86099; CAB06752.1; -.
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
KW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 22
FT CHAIN 23 904 GLYCOPROTEIN B.
FT DOMAIN 23 727 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 728 743 POTENTIAL.
FT TRANSMEM 749 768 POTENTIAL.
FT TRANSMEM 772 792 POTENTIAL.
FT DOMAIN 793 904 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 82 82 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 136 136 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 393 393 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 425 425 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 486 486 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 671 671 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 92 92 L -> V (IN REF. 1).
FT CONFLICT 198 198 T -> A (IN REF. 1).
FT CONFLICT 308 308 S -> T (IN REF. 1).
FT CONFLICT 438 438 L -> Q (IN REF. 1).
FT CONFLICT 568 568 S -> A (IN REF. 1).
FT CONFLICT 619 620 EL -> DV (IN REF. 1).
FT CONFLICT 636 636 R -> G (IN REF. 1).
SQ SEQUENCE 904 AA; 100217 MW; AB050A3AFB4F1066 CRC64;

Query Match 18.6%; Score 844; DB 1; Length 904;
Best Local Similarity 28.4%; Pred No. 3.1e-44;
Matches 247; Conservative 164; Mismatches 366; Indels 92; Gaps 27;

QY 33 TSPPTATWSTESPLTGHYTHDSSHGERRGNENRDEEQNNKIYGSPTFPYRVCSAG 92
Db 76 TTPPD---ANATVAAGHATLR--AHUREIKVENADAQ-----FVCPPT 115
QY 93 VGDVFRFQDTHVCPDASD-MVHSEGLLLYKQNIIFPMFRVRYKRVKVVITSTV- 144
Db 116 GATVQFEQPRRCPTREPGQNYTEGIAVFKENIAPYKFKATMYKDVTSVQWFGHYS 175
QY 145 -YNGIYSDSIITNQHTFYKSIPEWE--TEKMDTIYQCFNSLRLNTGNNLLTYVDRDINMT 201
Db 176 QFMGIFED-----RAPVPEEVIDKINTKVCSTAKYVNRNMETFAFRDHDHTD 226
QY 202 VFLQPVGDVTPDVKRYGSPQLYLEPGFWGYSYRRTTNCMLDMFARSNPFPFVTA 261
Db 227 MELKPAK-VATRTSRGWHHTDLKYNPSRVEAFHRYGTIVNCIVEVDARSVYDVFVLA 285
QY 262 TGDVTMSFPWS-GEDDH-ENKMKHEKPFWFSVINNY-KVVDYQNRGTVPGLKTRIFLDRE 318
Db 286 TGFVYMSFPFYGYREGSHTSHTSYAADRFKQVDFYARDLTTKARATSP--TTRNLLTTP 343
QY 319 EYTLSEKHLKMSYCPLTLKAFYNGIOTEHSGSVHFVANDITASFTT-----SKED 371
Db 344 KFTVAMDVWPKRPVACTMTKWQDEVNMLRAEYCGSFSSDAISTFTTNLTYSLSRVD 403
QY 372 MKEFNTHYHCLNBEIEKAEIEKKA-KVNSTHSHYGLKFKPTDGLLYVWQPLQIQLRLD 430
Db 404 LGD-----CIGRDAREADRMFARKYNATHIKVGQOYYLATGGFLIAYOPLLSNTLAE 457
QY 431 AKNKLANNETYSR-RSRQAEISTDPNMEM-TGNAGAGEYSSSENSITVAQOYAYDNLRI 488
Db 458 ----LVVREYMRQDRKPRNATPAPREAPSANASVERIKTSSIFARLQFTYNNHQRH 513
QY 489 INNILEDLSKAWCREQHRAALVWNLKSKINPTSVMSMIYNRPVSAKRIGDVISVNCIVV 548
Db 514 VNDMLGRIAVAWCELQNHLETLWNEARKLNPNNAIASATVGRVSARMLGDVMAVSTCVV 573
QY 549 DQTSVLSHLKSLRLSLSDEKCFSPVPVTFKFWNDSTIYKQGLGVNNELITTTTYLETCQE 608
Db 574 APDNVIVQNSMR-VSSRPGTCYSRPLVSFRYEDQGLIEGOLGNNELRLTRDALEPCTV 632
```


RESULT 20
VGLB HSV23

DB VGLB_HSV23 STANDARD; PRT; 904 AA.
P06763;
01-JAN-1988 (Rel. 06, Created)
01-JAN-1988 (Rel. 06, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B precursor.
GN GB OR UL27 OR GB2.
OS Herpes simplex virus (type 2 / strain 333).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10313;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87112925; PubMed=3027364;
RA Stuve L.L., Brown-Shimer S., Pacht C., Najarian R., Dina D.,
Burke R.L.;
RT "Structure and expression of the herpes simplex virus type 2
glycoprotein gb gene";
RL J. Virol. 61:326-335(1987).
CC -1- SUBUNIT: DIMER, PROBABLY LINKED BY DISULFIDE BONDS.
CC -1- MISCELLANEOUS: THERE ARE SEVEN EXTERNAL GLYCOPROTEINS IN HSV1 AND
2: GH, GC, GD, GI, AND GE.
CC -1- MISCELLANEOUS: GB IS THE ONLY GLYCOPROTEIN THAT IS KNOWN TO BE
REQUIRED FOR VIRAL GROWTH.
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; M15118; AAA45837.1; -
DR PIR; A26790; VGBEB2.
DR InterPro; IPR000234; Glycoprot B.
DR Pfam; PF00686; Glycoprotein B; 1.
DR ProDom; PD000693; Glycoprot B; 1.
KW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 904 GLYCOPROTEIN B.
FT DOMAIN 23 727 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 728 743 POTENTIAL.
FT TRANSMEM 749 768 POTENTIAL.
FT TRANSMEM 772 792 POTENTIAL.
FT DOMAIN 793 904 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 82 82 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 136 136 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 393 393 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 425 425 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 486 486 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 671 671 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 904 AA; 100186 MW; A8B36F74FDBC839 CRC64;
Query Match 18.5%; Score 839; DB 1; Length 904;
Best Local Similarity 28.3%; Pred. No. 6.4e-44;
Matches 246; Conservative 164; Mismatches 367; Indels 92; Gaps 27;
QY 33 TSPNTATWSTESPLTGHYTHDSHGCGNENRDSSEQNKNIYGSPTFPYRVCSASG 92
DB 76 TTPPD-----ANATVAAGHATLR--AHLREIKVENADAQ-----FVYCPPT 115
QY 93 VGDVFRFQTDHVCPSDAD-MVHSEGLLIYKQNIIPFMFRVRYKRVKVVTTSTV----- 144
DB 116 GATVVOEQPRCPTREGQNYTEGIAVFKENIAPYKATWYKDVTVSQVWFGRHYS 175
QY 145 -YNGIYSDSTNQHTFYKSTPEPE--TEKMDTYQCNSRLNTGGLNLIYVDRDDINMT 201
DB 176 QFMGIFED-----RAPVPEEVIDKINAKGVCRCSTAKYVRNNMETTAFHRDHDHTD 226
QY 202 VFQPVGDVTPDKRYGSOPELYLEPGWFGVSGYRRRTTVNCELMDMFARSNPFPDFVTA 261

DB 227 MELPKAK-VATRTSRGMHTTOLKYNPSRVEAFHRYGTTVNCVIEVDARSVYDFEFLA 285
QY 262 TGDVTVMSPFWS-GEDDH-ENKHEKPFVSVINNY-KVVDYQNRGTVPGLKTRIFLDRE 318
DB 286 TGDFTVMSPFYGYREGSHTHTSYAADRFKQVDGFYARDLTTTKARATSP--TTNNLLTTP 343
QY 319 EYTLSEKHLKMSYCPLTLWKAFYNGIOETHSGSYHFVANDITASFTT-----SKED 371
DB 344 KFTVANDWPKPAPVCTMTKQEVDEMLRAEYCGSFRFSDAISTFTTNLTQYSLSRVD 403
QY 372 MKEFNTTYHCLNEEIKAEIKKYA-KVNSTHSGYGLKFKTDGGLYLVQWQPLIQNRLLD 430
DB 404 LGD-----CIGRDAREADRMFAKYNATHIKVGQPYVATGGLIAYQPLLSNTLAE 457
QY 431 AKNKLNNETYSR-RSRQAEISTTDPMMEM-TGNAGAGEYSSSENSITVAQVQYADNLRIR 488
DB 458 ----LYVREYMRQDRKPRNATPAPREAPSANASVERIKTTSIEFARLQFTYNNIQRH 513
QY 489 INNILEDLSKACREOHRALVWNLKINPTSMYINRPVSAKRIGDIVSVNCIVV 548
DB 514 VNDMLGRIAVMCELOHBLTLWNEARKLNPNIAIATVGRVRSARMLGDVMAVSTCPV 573
QY 549 DOTSVLHSLRLLSASDEKCFSPPTVTFKFMNDSTIYKQGLGVNNEILLTITTYLETQOE 608
DB 574 APDNVIVQNSMR-VSSRPGTCYSRPLVSFRYEDQGPLIEGQLGNNELRLTRDALEPCTV 632
QY 609 NTEYFQAKTDMYIKNYEHLKTVPLSSITLDTFTALNFTLENVDFKVIELYITRDEKR 668
DB 633 GHRRYFIFGGYVYFPEYAYSHQLSRADVTTVSTFDLNTMLEDEHFEVPLEVYTRHEIK 692
QY 669 LSNVFDIETMFREYNYAQRVSLGRKDLDSLNRNQFVD--AFGLSMDDLG-AVGOTVV 725
DB 693 DSGLLDYTEVQRNQHLDRPADI-DTVIRADANAAMFAGLCAPFEGMGDLGRAVGKVM 751
QY 726 NAVSGVATLFSSIVTGFINFIKNPFGGMLMIIVIGLVFAIYFLTKTKIYETAPIKMIY 785
DB 752 GVVGWV----SAVSGVSFSMSNPFGLAVGLLVLAVAFPAFVYVQLQORNPWKALY 807
QY 786 PEIDKLKEREKSEIAPISEE-----ELERIVLAMHIHQONSHMETKTRKDP 832
DB 808 PLTTKELKTSDFGVGGEAGEGGAEGGDEAKLAEMIRYMYALVSAMERTEHKARKKG 867
QY 833 KDSIL-TRAQNM-LRKR--SGYSLNKNAE 857
DB 868 TSALLSSKVTNVLRKRNKARYSPLHNE 896
RESULT 21
VGLB_HSV1P STANDARD; PRT; 904 AA.
AC P08665;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B precursor.
GN GB OR UL27
OS Herpes simplex virus (type 1 / strain Patton).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10308;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87112925; PubMed=3027364;
RA Stuve L.L., Brown-Shimer S., Pacht C., Najarian R., Dina D.,
Burke R.L.;
RT "Structure and expression of the herpes simplex virus type 2
glycoprotein gb gene";
RL J. Virol. 61:326-335(1987).
CC -1- SUBUNIT: DIMER, PROBABLY LINKED BY DISULFIDE BONDS.
CC -1- MISCELLANEOUS: THERE ARE SEVEN EXTERNAL GLYCOPROTEINS IN HSV1: GH,
GB, GC, GD, GI, AND GE.
CC -1- MISCELLANEOUS: GB IS THE ONLY GLYCOPROTEIN THAT IS KNOWN TO BE
REQUIRED FOR VIRAL GROWTH.
CC

```

CC      -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL: K03541; AAA45778.1;
CC      InterPro: IPR000234; Glycoprot_B.
CC      Pfam: PF00606; Glycoprotein B; 1.
CC      ProDom: PD000693; Glycoprot_B; 1.
CC      Glycoprotein; Transmembrane; Signal.
CC      SIGNAL 1 30
CC      CHAIN 31 904
CC      DOMAIN 31 730
CC      TRANSMEM 731 746
CC      TRANSMEM 752 772
CC      TRANSMEM 775 795
CC      DOMAIN 796 904
CC      CYTOPLASMIC (POTENTIAL).
CC      N-LINKED (GLCNAC. .) (POTENTIAL).
CC      CARBOHYD 87 87
CC      CARBOHYD 141 141
CC      CARBOHYD 398 398
CC      CARBOHYD 430 430
CC      CARBOHYD 489 489
CC      CARBOHYD 674 674
CC      CARBOHYD 674 674
CC      SEQUENCE 904 AA; 100115 MW; 782581DC830A626F CRC64;
CC
CC      Query Match 18.4%; Score 836.5; DB 1; Length 904;
CC      Best Local Similarity 27.5%; Pred. No. 9.1e-44;
CC      Matches 244; Conservative 181; Mismatches 361; Indels 101; Gaps 28;
CC
Qy      24 SLSTAETGVTPPNTATWSTSP--LTGHYTHDSSHGE-RGNENRDSSEQNKNYIGSP 80
Db      58 ALGAAPTGDPKPKKVKPKNTPPPRPGDGNATVAAGHATLREHLDIKAEINTDANFY--- 114
Qy      81 STFPYRVCSAGVGDFRQFDHVCPCDASD-MVHSEGIILLYKONIIPFMRVRYKRVV 139
Db      115 -----VCPPTGATVVGFEQPRCPTRPEQNYTEGLAVVKEKNIAYKFKATMYKDV 168
Qy      140 TTSTV-----YNGIYSDSITNOHTFYKSIPEWE--TEKMDTIYQCFNSLRNTGGNL 189
Db      169 TVSQWVFGRHYSGQMGIFED-----RAPVPEEVIDKINAGVCRSTAKYVRNLE 219
Qy      190 LTYVDRDDINTVFLQPDGV-----TPDKYGSQPELYLEPGFWGSGYRRRTTVN 241
Db      220 TTAHRDDHETDMLKPNAAATRTSRGWHTTDLKYNPSRVEAF-----HRYGTTW 270
Qy      242 CELDMFARSNPFPDFVFTATGDTVMSPFWS-GEDDH-ENKMHEKPFVSVINNY-KVV 298
Db      271 CIVEVDARSVPYDFVFLATGDFVYMGFFYGYREGSHTHTSYAADRKFQVDGYFARDL 330
Qy      299 DYQNGTVPLGKTRIFLDREYTLISWEKHLKNMSYCLTLWKAFYNGIQTEHSGSYHFVA 358
Db      331 TTKARATAP--TTRNLLTPKFTVAWDVWPKEPSVCTWKQEVDEMELRSEYGSFRSS 388
Qy      359 NDITASFT-----SKEDMKFNTYTHCLNEEIKABIEKYYA-KVNSTHYSKYGDLKYF 410
Db      389 DAISTFTFTNTIETPLSRVDLGD-----CIGKDARDAMDRIFFARRYNATHIKVGQPOYY 442
Qy      411 KTDGGLYVWQPLIQRLLDAKNLNNETYRRSRROAESTTDPMMENTGNGAGEYSSE 470
Db      443 LANGGFLAYQLSNTLAE-----LYVREHUREQSRKPPNTPPPPGASANASVERIKTT 498
Qy      471 NSITVAQVQYAYDNIIRINNIEDLSKAWCREQRAALVWNLSEKINPTSVMSMYNRP 530
Db      499 SSIEFARLQFTYNIHQHVNDMLGRVATWCELQNHLETLNNEARKLNPNIAASATVGR 558
Qy      531 VSAKRIGDVISVNCIVVDQTSVSLHKSRLLSASDEKCFSRPPVTFKFMNDSTIYKQL 590
Db      559 VSARMLGDVMAVSTCVPAADNVIVQNSMR-ISSRPGACYSRPLVSPRYEQDGLVEQL 617

```

```

Qy      591 GWNNEILLTITTYLETCOENTYFYFOAKTDMVIYKYNHKLKTVPLSSITLTLDTFIALNFTL 650
Db      618 GENNELRLTRAIPCTVGHRRYTFGGGYVYFESAYSHQLSRADITTVSTFIDLNTM 677
Qy      651 LENVDKVIETYRDEKRLSNVFIETMFRFYNYYAQKRVSLRDLDLSTNRNOF--VD 708
Db      678 LEDHEFVPLEVYTRHEIKDGLDYTEVQRNQLHDLRFADI-DTVIHADANAAMFAGLG 736
Qy      709 AFGSLMDDLG-AVGQTVVNAVSGVATLFPSSIVTGSFINKPNPFGMLMIIVIGVLFAIY 767
Db      737 AFFEGMGDLGRAVGKVMGIVGGVV-----SAVSGVSFMSNPPFGALAVGLLVLAGLAAAF 792
Qy      768 FLTKTKIYETAPIKMIYPEIDK-LK-----ERECKSEIAPISE-EELERIVLA 814
Db      793 FAFYVMRLQSNPKALYPLTTKELKNTPDASGEGEEDGDFDEAKLAAREMIRYMAL 852
Qy      815 MHIHQQNSHMETKTRKPKOSILT-RAQNML---RKRSGYSNLKNAE 857
Db      853 VSAMERTEH---KAKKKGTSALLSAKVTDMMVRKRRNTNYTQVPNKD 896
CC
CC      RESULT 22
CC      VGLB ILTVS STANDARD; PRT; 883 AA.
CC      ID VGLB ILTVS
CC      AC P27415;
CC      DT 01-AUG-1992 (Rel. 23, Created)
CC      DT 01-AUG-1992 (Rel. 23, Last sequence update)
CC      DT 16-OCT-2001 (Rel. 40, Last annotation update)
CC      DE Glycoprotein B precursor.
CC      GN GB.
CC      OS Infectious laryngotracheitis virus (strain SA-2) (ILTV).
CC      OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
CC      OC Alphaherpesvirinae; Infectious laryngotracheitis-like viruses.
CC      OX NCBI_TaxID=10343;
CC      RN [1]
CC      RP SEQUENCE FROM N.A.
CC      RX MEDLINE=91335774; PubMed=1840710;
CC      RA Kongsuwan K., Prideaux C.T., Johnson M.A., Sheppard M., Fahey K.J.;
CC      RT "Nucleotide sequence of the gene encoding infectious
CC      RT laryngotracheitis virus glycoprotein B.";
CC      RL Virology 184:404-410 (1991).
CC      CC -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL: M64927; AAA88009.1; ALT SEQ.
CC      PIR: A40567; VGBEIS.
CC      DR InterPro: IPR000234; Glycoprot_B.
CC      DR Pfam: PF00606; Glycoprotein B; 1.
CC      DR ProDom: PD000693; Glycoprot_B; 1.
CC      KW Glycoprotein; Transmembrane; Signal.
CC      SIGNAL 1 26
CC      CHAIN 27 883
CC      DOMAIN 27 724
CC      TRANSMEM 725 744
CC      TRANSMEM 751 771
CC      DOMAIN 772 883
CC      CARBOHYD 102 102
CC      CARBOHYD 121 121
CC      CARBOHYD 211 211
CC      CARBOHYD 262 262
CC      CARBOHYD 360 360
CC      CARBOHYD 579 579
CC      CARBOHYD 635 635
CC      CARBOHYD 649 649
CC      SEQUENCE 883 AA; 100142 MW; B414083346C55014 CRC64;

```

Query Match 18.4%; Score 836; DB 1; Length 883;
 Best Local Similarity 27.9%; Pred. No. 9.4e-44;
 Matches 222; Conservative 158; Mismatches 308; Indels 108; Gaps 23;

QY 57 SHGERGNNRND--SEBQNKIY-----GSPSTFPYR-----VCSASGVGDVFRFQFDHV 104
 DB 34 SHGIGWNSPHDTASMDVGKISFSEAGISGAPKEQIRNRIACSSPTGASVARLAQPRH 93
 QY 105 CPDASDMVH-SEGILLLYKQNIIFPMFRVRKYRKVTTSTVYNGIYSDSTNQHTFYKSI 163
 DB 94 CHRHADSTNMTAGIAVFKQNIAPVFNVTLYKHITTTTWTWALFSRPOITNEYVTRVPI 153
 QY 164 EPWETEKMDITIOCFNSLRNTGNNLTYYVDRODINMTVLOP-----VDGVTPDV 214
 DB 154 DYHEIVRIDRSGECSSKATYHKMFNFEEAVDNDAEKKLPVPSLLRSTVSKAPHITNT 213
 QY 215 KRYGQPELYLEPGWFGWYRRRTTVCNCELMDFMARSNPPDFEVTATGDTVMSPFWSG 274
 DB 214 KRHQTL-----GYRTSTSDVCVVEYLQARSVPYDYEGMATGDTVELSPFYT- 260
 QY 275 EDDHENKHEKPFVSVINNYK---VVDYQNR-----GTVPLGKTRIFLDREYVTLSEKH 327
 DB 261 -----KNWTPGRHRSYRDYRFLFIANYQVRDLETGQIRPPKKRNFELTDEQFTIGWDM 314
 QY 328 LKMSYCPCLTLKAFYNGIQTEHSGSYHFVANDITASFTTSKEDMKEFNTT-----YHCL- 382
 DB 315 EEKESVCTLKWIEVPAVRVSNYSYHFLSKDMTMTFSSGKQ---PFIISRLHLAECVP 371
 QY 383 ---NEEKAEIEKKYAKVNSTHYSKGLKYPKPTDGLYLVNQPLIQNRLLD---AKNKLN 436
 DB 372 TIASEADIGIPARKY---SSTHVRSGLIEYVYLGSGGLIAFQKLSHGLAEMLYEEAQRQ 428
 QY 437 NETYSRRSQAESTTDPMMEMTNGAGGEYSSENSITVAQVQVAYNLRINRIELEDL 496
 DB 429 NHPGRERRQAQRRTASLQ--SGPQGDRIITHTSSATFAMQLPAYDKIOAHVNEILGNL 486
 QY 497 SKACRQHQRAALVWNLKINPTSVMSLYNRPVSASAKRGDVISVNSCTVVDQTSVSLH 556
 DB 487 LEAWCEIQNLQIVHHEKLNPNLSMTSLFGQVPSARLLGDIVAVSKCIEPIENIRMQ 546
 QY 557 KSLRLLSASDEKFSRPPVTFKFM-----NDSTIYKQGLGVNNEILLTYYL 603
 DB 547 DSMR-VGDPDTCYTRVPLIFRYSSSPESQFSANSTENHNLGILQGLGEHNEILQGRNLI 605
 QY 604 ETCOENTYYPQAKDMMYIKVNEHLKTVPLSSITLDTFIALNFTLLENVDFKVIELYT 663
 DB 606 EPCMINHRRYFLGENTYLLYEDYTFVRQVNASIEEVSFTINLNAITLEDLDFVPVEVYT 665
 QY 664 RDEKRLSNVFDIETMREYNVYQAR-----VSGLRKDLDDLSTNRNQFVDAFGSLMDD 716
 DB 666 REELRDTGTLYNDVVRQYQNIYKRFIDITVIRGDRGDAI-----PRAIADFGNLTGE 720
 QY 717 LG-AVGQTVNVNAVSGVATLFSSIVTGFINKPFGQM-----LMIITVIGVLFAIYFLT 770
 DB 721 VGRALGTVMVTAATAAVI-----STVSGIASFLSNPFAALAGIAVAVVSIILGLLAFKYMVN 776
 QY 771 KTKIYETAIKMIYP 786
 DB 777 LKSN-----PVQVLIFF 787

RESULT 23
 VGLB_HSVB
 ID_VGLB_HSVB STANDARD; PRT; 980 AA.
 AC P28922;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glycoprotein B precursor (Glycoprotein 14).
 GN GB OR GP14 OR 33.
 OS Equine herpesvirus type 1 (strain Ab4p) (EHV-1), and
 OS Equine herpesvirus type 1 (strain Kentucky D) (EHV-1).

OC Viruses: dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.
 OX NCBI_TaxID=31520, 10330;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AB4P;
 RX MEDLINE=92295566; PubMed=1318606;
 RA Telford E.A.R., Watson M.S., McBride K., Davison A.J.;
 RT "The DNA sequence of equine herpesvirus-1.";
 RL Virology 189:304-316(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Kentucky D;
 RX MEDLINE=90219246; PubMed=2157895;
 RA Guo P.X., Goebel S., Ferkus M.E., Taylor J., Norton E., Allen G.,
 RA Languet B., Demetree P., Paoletti E.;
 RT "Coexpression by vaccinia virus recombinants of equine herpesvirus 1
 RT Glycoproteins gp13 and gp14 results in potentiated immunity.";
 RL J. Virol. 64:2399-2406(1990).
 CC -/- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: M86664; AAB02468.1; -;
 DR EMBL: M35145; AAA46067.1; -;
 DR PIR: G36798; VGBEG6.
 DR InterPro: IPR000234; Glycoprot_B.
 DR Pfam: PF00606; Glycoprotein_B; 1.
 DR ProDom: PD00693; Glycoprot_B; 1.
 DR Glycoprotein; Transmembrane; Signal.
 FT SIGNAL 1 86 POTENTIAL.
 FT CHAIN 87 980 GLYCOPROTEIN B.
 FT DOMAIN 87 852 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 853 870 POTENTIAL.
 FT DOMAIN 871 980 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 165 165 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 275 275 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 380 380 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 423 423 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 497 497 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 514 514 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 515 515 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 560 560 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 727 727 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 749 749 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 980 AA; 109805 MW; A6DDFA8CA5550FF5 CRC64;

Query Match 18.4%; Score 835; DB 1; Length 980;
 Best Local Similarity 26.9%; Pred. No. 1.3e-43;
 Matches 250; Conservative 161; Mismatches 383; Indels 134; Gaps 31;

QY 32 VTSPPNTATWSTESPLTGHYTHDSHSGRGNENRDSBQNKNIYGSPSTFPYRVC-SA 90
 DB 83 VRAVPTTSPPTSTPTS--MSTH--SHG-----TVDPTLLPTTDPDLRLAVRE 127
 QY 91 SGV-----GD-----VRFQTDHVCDDAS-DMVHSEGLILLYKQNIIFPMFRVRK 134
 DB 128 SGLAEDGDFYTCPPPTGTVVRIEPPRTCPKFDLGRNFTGIAVFKENIAPYKFRANV 187
 QY 135 YRKVVTTSTVYNGIYSDSITNQHTFYKSIETPETER-----MDTIYOCFNS---LRINTGG 187
 DB 188 YKDIIVTVRWKGYSHTSLSDR---YNDRVPSVSEIFGLIDSKGKSSKAEVLR-----D 240
 QY 188 NLLTYV---DRDDINMTVFLQPDVGVTPDKVRYGSGQPELYLEPGWF-WGSRRTTNC 243
 DB 241 NIMHAYHDEDEVELD--LVPSKATPGARAWQTNDTTSYVGMWPMRHY-TSTSVNCI 297

Db 500 RHVNDMLGRITATWCEQLQNRRLTWNERRLNPALASATVGRVRGARMGLDVMVASTCV 559
 QY 547 VVQTSVSLHKSRLLSASDEKCFSPPTVTFKFWNDSTIYKGLGVNNEILLTTTILETC 606
 Db 560 PVAPDNVIMONSIG-VAAREPGTCYSRPLVSFRYEADGPLVEGQGDNEIRLERDALEPC 618
 QY 607 QENTEYVFOAKTDMYIKYNEHLKTVPLSSITITLDTFIALNFTLLENVDFKVIETLTRDE 666
 Db 619 TVGHRRYFTGAGVYVEEYAYSHQLGRADVTTVSTVFINLNLWLEDFEVPLEVTRQE 678
 QY 667 KRLSNVFDIETMFREYNYYAQRVSGRLKOLLDSLTRNQFVDAFGSLMDLGAVGQTVVN 726
 Db 679 IKDSGLLDYEVQRNQLHALRFADI-DTVIKADAAHALFAGLY-SFFEGLGDVGRAVGK 736
 QY 727 AVSGVATLFSISIVTGFNFNFKPFGMLMIIVIGVLFAIYFLTKTKIYETAPIKIYIP 786
 Db 737 VVNGIVGGVVSAYSGVSSFLSNPFGALVGLLVLGALAAFFAFRYVMRLQRNPKALYP 796
 QY 787 EIDKLKERECKSIAPIS-----EELERIVLA-----MHIHQONSHMETKTRKDP 832
 Db 797 ----LTKELKSGAPLAGGEGDEGAEFDFAKLAQAREMIRYMALVSAMERTEHKARKG 852
 QY 833 KDSIL-----TRAQNMLRKRGYGNLKNASVEM 861
 Db 853 TSALLSAKVTDAVMRKARPRYSPLRDTDEEL 885

RESULT 25
 VGLB_VZVD STANDARD; PRT; 868 AA.
 AC P09257;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glycoprotein B precursor (Glycoprotein II).
 GN 31.
 OS Varicella-zoster virus (strain Dumas) (VZV).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.
 OX NCBI_TaxID=10338;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86306657; PubMed=3018124;
 RA Davison A. J., Scott J. E.;
 RT "The complete DNA sequence of varicella-zoster virus.";
 RL J. Gen. Virol. 67:1759-1816(1986).
 CC -/- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X04370; CAA27914.1; -;
 DR PIR; E27214; VGBE31.
 DR InterPro; IPR00234; Glycoprot B.
 DR Pfam; PF00606; Glycoprotein_B; 1.
 DR ProDom; PD000693; Glycoprot_B; 1.
 KW Glycoprotein; Transmembrane; Signal.
 FT SIGNAL 1 ?
 FT CHAIN ? 868 GLYCOPROTEIN B.
 FT DOMAIN ? 679 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 680 695 POTENTIAL.
 FT TRANSMEM 701 720 POTENTIAL.
 FT TRANSMEM 724 744 POTENTIAL.
 FT DOMAIN 745 868 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 84 84 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 194 194 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 372 372 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 440 440 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 557 557 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 623 623 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 868 AA; 98066 MW; 5898E1346F4FF902 CRC64;
 Query Match 18.3%; Score 832; DB 1; Length 868;
 Best Local Similarity 26.7%; Pred. NO. 1.6e-43;
 Matches 236; Conservative 168; Mismatches 365; Indels 116; Gaps 27;
 QY 40 TWSTESPLTGHVTHDSSHGERNNENRDSSEQNKNIIYGSPTFFPYRVCSASGVDFRF 99
 Db 23 TOSIEDITRSAHLGDGD-----EIRAIHKSQDAETKPT---FYVCPPTGSTIVRL 70
 QY 100 QTDHVCDAASDMVH-----SEGILLIYKQNIIPFMRVRKRYKRVVTTSTVNGIYSDSIT 154
 Db 71 ETRICPD---YHLGKNFTGIAVYVYKNIAYKFAIYVYKDVIVSTAWAGSSYQIT 126
 QY 155 NQHTFYKSIETPWE-TEKMDTIYOCFNSRLNTGNNLLTV-----DRDDINMTV 202
 Db 127 NRYADRVPIPVSEITDTIDKFGKC-----SSKATVVRNNHKVFAFNEDKNPQDMPL 177
 QY 203 FLOPVDGVTVDVKRYGSOPELYLEPCWFWGSVYRRRTVNCELMDMFARSNPPDFVTAT 262
 Db 178 IASKYNSV--GSKAMHTTNDTYMVAG--TPTGYRTGTSVNCIIEVEARSIFPYDSFGLST 234
 QY 263 GOTVEMSPWSCGD-----DHEN-----KMHEKWFVSVINNYKVVDYQNRGTVPGLKTRIF 314
 Db 235 GLIIMSFPFGLRDGAYRHSNYAMDRFHQ-----PEGYRQRLDLDTRALLE-PAARNF 286
 QY 315 LDREEYTLSEKHLKNMSYCLTLWKAFYNGIQTEHSGSYHFVANDITASFTTSKEDMKE 374
 Db 287 LVTPHLTVGMWPKRTEVCSLVKVRVEDVVRDEYAHNFRFTMTKLTSTTFIS---ETNE 343
 QY 375 FNTTY-----HCLNEELKABIEKKY-AKVNSTHSKYGDLYKFKTDGGLYLWQBLIQRLL 429
 Db 344 FNLNLQIHLSCQVKEEARAIINRIYTRYNSHVRTGDIQTYLARGGVVVFQPLLNSLA 403
 QY 430 ----DANKNLNNTYRSRROAESTTDPMMEMTGNAGGAYESSSENSITVAQVQAYDN 484
 Db 404 RLYLQELVRENTNHSPOKHPTRNRSRSPVELRANRT---ITTTSSVEFAMLQTYDH 460
 QY 485 LRIRINNILEDLSKAWCREQHRALVWNLSEKINPTSVMSIYNRPVSAKRIGDVISVN 544
 Db 461 IQEHVNMELARISSSWCQLQNRERALSGLFPINPSALASTILDQRVKARILGDVISVN 520
 QY 545 CIWV-DQTSVSLHKSRLLSASDEKCFSPPTVTFKFWNDSTIYKGLGVNNEILLTTTIL 603
 Db 521 CPGLGSDTRIILQNSMR-VSGSTTRCYRPLSIVSLNGSGTVEGQLGTDNELIMSRDLL 579
 QY 604 ETCQENTYVFOAKTDMYIKYNEHLKTVPLSSITITLDTFIALNFTLLENVDFKVIETLYT 663
 Db 580 EPCVANHKRYFLFGHHYVYEDYRYVRETAHVDVGMISTYVDNLNLTLLKDRFEMPLOVYT 639
 QY 664 RDEKLSNVFDIETMFREYNYYAQRVSGRLKDL-LDLSTNRNQFVDAFGSLMDLGAVGQ 722
 Db 640 RDELRTDGLDYSEIQRRNQMSHLRFYDIDKVVQYDSGT---AIMQGMQAFQFGLGTAGQ 696
 QY 723 TVNVAVSGVATLFSISIVTGFNFNFKPFGMLMIIVIGVLFAIYFLTKTKIYETAPIK 782
 Db 697 AVGHVVLGATGALLSTVHGFTTFLSNPFGALVGLLVLGALAAFFAFRYVLKLTSPMK 756
 QY 783 MLYPEIDK-LKE-REGKSEIA-----PISE-----EELERIV 812
 Db 757 ALYPLTTKGLKQLPEGMDPFAEKPNATDTPIEIGDSQNTSPSVNSGFPDPDKFREAEQMI 816
 QY 813 LAMWHIHOQNSHMETKTRKDPKDSILTRAQ---NMLKRSGYGNL 854
 Db 817 KYMTLSAAERQESKARKKNTKTSALLTSRLTGLALNRGRYSRVR 861

RESULT 26
 VGLB_ILTVT
 ID VGLB_ILTVT
 AC P24904;
 STANDARD; PRT; 883 AA.

DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glycoprotein B precursor.
 GN GB.
 OS Infectious laryngotracheitis virus (strain Thorne V882) (ILTUV).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Infectious laryngotracheitis-like viruses.
 OX NCBI_TaxID=10344;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91132136; PubMed=1847176;
 RA Griffin A.M.;
 RT "The nucleotide sequence of the glycoprotein gB gene of infectious laryngotracheitis virus: analysis and evolutionary relationship to the homologous gene from other herpesviruses."
 RL J. Gen. Virol. 72:393-398 (1991).
 CC -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D00818; BAA00699.1; --
 DR PIR; A38478; VBELL.
 DR InterPro; IPR000234; Glycoprot_B.
 DR Pfam; PF00606; Glycoprotein_B; 1.
 DR ProDom; PD000693; Glycoprot_B; 1.
 KW Signal; Glycoprotein; Transmembrane.
 FT SIGNAL 1 26 POTENTIAL.
 FT CHAIN 27 883 GLYCOPROTEIN B.
 FT DOMAIN 27 724 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 725 744 POTENTIAL.
 FT TRANSMEM 751 771 POTENTIAL.
 FT DOMAIN 772 865 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 102 102 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 121 121 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 211 211 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 262 262 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 360 360 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 579 579 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 635 635 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 649 649 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT SEQUENCE 883 AA; 100160 MW; D457DF5178840A7D CRC64;
 Query Match 18.3%; Score 830; DB 1; Length 883;
 Best Local Similarity 27.8%; Pred. No. 2.2e-43;
 Matches 221; Conservative 159; Mismatches 308; Indels 108; Gaps 23;
 QY 57 SHGERGNENRD--SEEQKNY-----GSPSTPPYR----VCSASGVGVFRFQTDHV 104
 DB 34 SHGIAGIDPRDASDVGKISFSAIGSGAPKEQIQRNIRPACSPPTGASVARLAQPRH 93
 QY 105 CPDASDMVH-SGILLIYKQNIIPMFRVKRKVVTTSTVNGIYSIDTNGHTFYKSI 163
 DB 94 CHRHADSTNMTGIAVVFVKQNIAPVFNVTLYKHITTTWALFSPQITNEYVTRVPI 153
 QY 164 EPMETEKMDTIYQCNLSRLNTGGNLLTVVDRDDINMTVFLOP-----VDGVTPDV 214
 DB 154 DYHEIVRDSRGECSKATYHKNFMMFFEAYDNDEAKKLPLVPSLLRSTVSKAFHTTFT 213
 QY 215 KRYGQPELYLEPGFWGWSYRRRTTNCLEMDMFARSNPPFFFTATGDTVEMSPFWSG 274
 DB 214 KRHTL-----GYRTSTVDCVVEVLQARSVPDYFGWAGDVTVEISFFYT- 260
 QY 275 EDDHKNKHEKFWFVSINNYK---VDVQNR-----GTVPGLGKTRIFLDREYTLWSKX 327
 DB 261 -----KNTTGPRRHSVYRDYRFLFIANYQVRDLETGTQIRPPKKNFLTDQFTIGWDAM 314

QY 328 LKNMSYCPCLTLWKAFYNGIQTEHSGSYHFVANDITASFTTKEDMKERNTT----YHCL- 382
 DB 315 EEKESVCTLSKWEIEVPEAVRVSYKNSYHFSKDMTMTFSSGKQ---PNISRLHLAECVP 371
 QY 383 ---NEEIKAEIEKYYAKVNSTHSHYKDYKPKTDGGLYLVWQPLIQNRLD---AKNKLN 436
 DB 372 TIASEAIDGIFARKY---SSTHVRSGDIEYVLGGGFLIAFQKLMHGLAEWYLEEAQRQ 428
 QY 437 NETYSRRSRROAESTTDPMMEMTGNGAGGEYSSENSITVAQVQYAYDNLRIINNILEDL 496
 DB 429 NHLPRGRERQAQRRTASLQ--SGPOGDRITTHSSATFAMLOFAYDKIOAHVNELIGNL 486
 QY 497 SKAWCREQHRRAALVMNELSKINPTSVMSIMYNRPVSAKRIGDVISVNCIVVDQTSVSLH 556
 DB 487 LEANCELQNRQLIYVHHEMKLNPNLSLTSIFGQPSARLLGDIVAVSKCIPIENIRMQ 546
 QY 557 KSLRLLSASDBKCFSRPPVTFKFM-----NDSTIYKQGLGVNNEILLTTLTYL 603
 DB 547 DSMR-MPGDPTWCYTRPVLIFRYSPPSQFSANSTENHNLDILQGLGEHNEILQGRNLI 605
 QY 604 ETCQENTYFYQAKTDMYIKNYEHLKTVPLSSITTLDTFFIALNFTLLENVDFKVIELTY 663
 DB 606 EPCMINHRRYFLGELNYLLYEDYTFVRQVNASEIEVSIFINLNATILEDLDFVPEVYT 665
 QY 664 RDEKRLSNVFDIETMFRYNYAQR-----VSGLRKDLDLSTNRNQFVDAGSLMDD 716
 DB 666 REEURDTGTLNYDDVVRQNIYKRFDRIDTVIRGDRGDAI-----FRAIADFFGNTLGE 720
 QY 717 LG-AVGQTVNNAVSGVATLFSISIVTGFINFKNPGGM-----LMIIVVIGLVFAIYFLT 770
 DB 721 VGKALGTVMVMTAAAAVI---STVSGIASFLSNFPAALGIGIIVVSVIILGLLAFKYVMN 776
 QY 771 KTKIYETAPIKMIYP 786
 DB 777 LKSN-----PVQVLFP 787
 RESULT 27
 VGLB ILTV6 STANDARD; PRT; 873 AA.
 AC Q02409;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glycoprotein B precursor.
 GN GB.
 OS Infectious laryngotracheitis virus (strain 632) (ILTUV).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Infectious laryngotracheitis-like viruses.
 OX NCBI_TaxID=31521;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92180317; PubMed=1665614;
 RA Poulsen D.J., Adams Burton C.R., O'Brian J.J., Rabin S.J.,
 RA Keeler C.L. Jr.;
 RT "Identification of the infectious laryngotracheitis virus glycoprotein gB gene by the polymerase chain reaction."
 RL Virus Genes 5:335-347 (1991).
 RN [2]
 RP REVISIONS.
 RA Keeler C.L. Jr.;
 RL Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X56093; CAA39573.1; --

```

DR PIR; S26690; S26690.
DR InterPro; IPR000234; Glycoprotein B.
DR Pfam; PF00606; Glycoprotein B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
KW Signal; Glycoprotein; Transmembrane.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 873 GLYCOPROTEIN B.
FT DOMAIN 22 714 EXTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 715 734 BY SIMILARITY.
FT TRANSMEM 741 761 BY SIMILARITY.
FT DOMAIN 762 855 CYTOPLASMIC (BY SIMILARITY).
FT CARBOHYD 92 92 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 111 111 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 201 201 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 252 252 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 350 350 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 569 569 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 625 625 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 639 639 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 873 AA; 99039 MW; E73E7F6D7A0E3AAD CRC64;

Query Match 18.2%; Score 829; DB 1; Length 873;
Best Local Similarity 27.8%; Pred. No. 2.5e-43;
Matches 221; Conservative 158; Mismatches 309; Indels 108; Gaps 23;

QY 57 SHGERGNEIRD--SEQNKNIV-----GSPSTFPYR---VCSAGVGDFVRFQDTHV 104
DB 24 SHGIAGIDPDTASMDVGKISFEAIGSGAPKEPQIRNIFACSSPTGASVARLAQPRH 83

QY 105 CPDASDMVH-SEGILLIYKQNIIPFMFRVRYKRVVTTSTVYNGIYSDSITNQTFFYKSI 163
DB 84 CHRHADSTNMTGIAVVFVKQNIAPVFNVTLYKHITVTITWALFSRQITNEVYTRVPI 143

QY 164 EPWTEKMDTYYQCFNSLRNLGTGNTLYVDRRDINDMTVFLOP-----VDGVTDPV 214
DB 144 DYHEIVRIDRSGECSSKATYHKNFMFEAYDNDAEKKLPLVPSLLRSTVSKAFHTTFT 203

QY 215 KRYGQPELYLEPGWFGSVYRRRTTVCCLMDMFARSNPPDRFVTATGDTVEMSPWG 274
DB 204 KRHOTL-----GYRTSTVDCVBYEQARSVPYDYGMATGDTVEISFPYT- 250

QY 275 EDDHENKMKHPFVSVINNYK---VVDYQNR-----GTVPGLKTRIFLDREYTLSEKH 327
DB 251 -----KNTTGPRRHSVVDRYFLEIANQYVRDLETGQIRPPKKRNFELTDEQFTIGWAM 304

QY 328 LKWSYCELTLLWAFYNGIQTEHSGSHFVANDITASFTSKEDMKEFNTT-----YHCL- 382
DB 305 EEKESVCTLSKWIEVPEAVRVSVKSYHFSLKDMTMTFFSGKQ---PFNISRHLHAECPV 361

QY 383 ---NEETKAEIEKKYAKVNSHYSKYGDLYKPKTDGGLYLVWQPLIQNRLLD---AKNKLN 436
DB 362 TIATEADIGTFARKY---SSTHVRSGDIEYYLGGGFLIAFOKUMSHGLAMEYLEEAQRQ 418

QY 437 NETYSRRSRROAESTTDPMMEMTNGAGGEYSSENSITVAQVQVAYONLRIRINILLEDL 496
DB 419 NHPGRGRERQOACGRRTASLQ--SGPQGDRTTHSSATFAMLOPAYDKIOAHVNELIGNL 476

QY 497 SKAWCRQHRALVWNLKINPTSVMSMYNRPVSAKRIGDVISVNCIVVQDTSVSLH 556
DB 477 LEAWCELQNRQLIYVHMKMLNPNLSMTSLFGQVPSARLLGDIYAVASKCIEPIENIRMQ 536

QY 557 KSLRLLSASDEKCFSPRPVTFKFM-----NDSTIYKQGLGVNNEILLTITL 603
DB 537 DSMR-MFGDPTMCTYRPLVIFRYSSESPESFANSNTENHNDLILGQGEHNEILQGRNLI 595

QY 604 ETCQENTYFQAKTDMYIYKNYEHKLTVPFLSSITTLDTFIALNFTLLENVDFKVIPLYT 663
DB 596 EPCMINHRRYFLLGENVLLYEDYTFVRQVNASBIEEVSIFINLNATILEDLDLDFVPVEVT 655

QY 664 RDEKRLNVDFIETMFREYNYAQR-----VSLGRKDLDDLSTNRNQFVDFAGSLMDD 716
DB 656 REELRDTGTLNDDVRYQNIYKRRFRDIDTIVRGDRGDAI-----FRAIADFPNGTLGE 710

```

```

QY 717 LG-AVGQTVNAVSGVATLFSSIVTGFINFIRKPFQGM-----LMIIVVIGLVFAIYFLT 770
DB 711 VGRALGTVMVMTAAAVI-----STVSGIASFLSPFAALGIGIAVWVSIIILGLLAFKVMN 766

QY 771 KTKKIYETAPIKMIYP 786
DB 767 LKSN-----PVQVLPF 777

RESULT 28
VGLB HSVE1 STANDARD; PRT; 980 AA.
AC PI8050;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B precursor (Glycoprotein 14).
GN GB OR GP14 OR 33.
OS Equine herpesvirus type 1 (isolate HVS25A) (EHV-1).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10327;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89279217; PubMed=2543744;
RA Whalley J.M., Robertson G.R., Scott N.A., Hudson G.C., Bell C.W.,
RA Woodworth L.M.;
RT "Identification and nucleotide sequence of a gene in equine
RT herpesvirus 1 analogous to the herpes simplex virus gene encoding the
RT major envelope glycoprotein gB.";
RL J. Gen. Virol. 70:383-394(1989).
CC - SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D00401; BAA00304.1; ALT_SEQ.
DR PIR; A31241; VGBE2H.
DR InterPro; IPR000234; Glycoprot B.
DR Pfam; PF00606; Glycoprotein B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
KW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 86 POTENTIAL.
FT CHAIN 87 980 GLYCOPROTEIN B.
FT DOMAIN 87 852 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 853 870 POTENTIAL.
FT DOMAIN 871 980 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 165 165 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 275 275 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 380 380 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 423 423 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 497 497 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 514 514 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 515 515 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 560 560 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 727 727 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 749 749 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 980 AA; 109931 MW; 94B69AF882FB6AC2 CRC64;

Query Match 18.1%; Score 824; DB 1; Length 980;
Best Local Similarity 26.9%; Pred. No. 5.9e-43;
Matches 252; Conservative 155; Mismatches 379; Indels 150; Gaps 31;

QY 32 VTSPPNATWSTESPLTGHGTHDSSHGGERGNNEEDSEQNKNIYGSPTFPVRYC-SA 90
DB 83 VRAVPTSPPTSTPTS--MSTH--SHG-----TVDPTELLPTETPDPLRAVRE 127

QY 91 SGV-----GD-----VFRFQDTHVCPDAS-DMVHSEGLIYKQNIIPFMFRVK 134

```



```

Db 128 SGILAEADGDFYTCPPPTGTVVRIPPTCPKFDLGRNFTGIAVIFKENTAPIYKFRANV 187
Qy 135 YRKVVTSTVYNGIYSDSTNQHTFYKSIPEWETEK-----MDTIYQCNSLRNLNTGGNLL 190
Db 188 YKDIWTVRWKGYSHTSLSDR--YNDRPVSVSVEEIFGLIDSKGKSKAB-----236
Qy 191 TYVDRDDINMTVFLQGVDPDVKRYGSGQPELYLE-----PGWF-WGSYR 235
Db 237 -YL-RDNIMHHAYHDEDEVELDLCR---PSQLRGARAWQTTNDTTSYVGMPWRHY- 289
Qy 236 RATTVNCELMDMFARSNPFPFVATGDTVMSPFW-----SGEDDHENKMKHEKPMFVSV 291
Db 290 TSTSYNCIVEEARSVPYDSFALSTGDIVVASPFYGLRAAARIEHNSYAQER---FRQ 346
Qy 292 INNYKVVDPYQNGRTVPLGKTRIFLDREEVTLSEWKLKNMSYCPPLTLAKAFYNGIQTEHS 351
Db 347 VEGYPRDLDSKLQAEPPVTKNFITTPHVTVSNWNTTEKKVEACTLTWKVEDELVRDEFR 406
Qy 352 GSYHFVANDITASFTTSKEDMK-BENTTYHCLNEBEIKAEIEKKYAK-VNSTHSHKYGDLKY 409
Db 407 GSYRTIRSIISSTFISNTTQFKLESAPLTCVCVSKAEKAIDSIIYKKQYESTHVSFSDVEY 466
Qy 410 FKTDCGLYLWQPLI-----QNRLDAKNKL-----NNETYSSRRRQAEST 451
Db 467 YLARGGLIAFRPMLSNELARLYLNLVRSNRTYDLKNLLPNANNNNTTTRRRSLLSV 526
Qy 452 TDPMMEMTG-----NGAGEVSS-----NITVAQVQYAYDN 484
Db 527 PEPOQTQGVHREQLHLRHKRAVATAGTSSNVTAKOLELIKTTSEIFAMQLQFAYDH 586
Qy 485 LRIRINILEDLSKAWCEQRAALVWNELSKINPTSVMSIYNNRPVSAKIGDIVSYSN 544
Db 587 IQSHVNEMLSRATACWTQNKERTLWNEWKINPISATLDERVAARVLGDVIAITH 646
Qy 545 CIVDQTSVSHKURLLSASDEKCFRPPVTF---KFMNDSTIYKGLGVNNEILLTTT 601
Db 647 CAKI-EGNVYLQNSMR---SMDSNTCYSRPPVTFITTKNANNRSGTEGQGEENEIFTERK 703
Qy 602 YLETCQENTYFQAKTDMYIKNYEHLKTVPLSSITTLDTFIALNFTLLENVDFKVEL 661
Db 704 LIEPCALNOKRYFKGKGYVYENYTFVRKVPPTIEIVSYVNLNLTLLDREFLPLEV 763
Qy 662 YTRDEKRLSNVFDIETMFREYNYAQRVSGLRKLLDLSLNRNQVDFAGSLMDLGLGAVG 721
Db 764 YTRAELEDTGLDISEIQRNQHLALREYDI--DSVNVNDVTAVIMQGIASFFKGLGKVG 821
Qy 722 QTVNAVSGVATLFSSIVTGFINFKNPFGLMLIIVVIGLFAIYFLTKTKIYETAPI 781
Db 822 EAVGLVLGAAGAVVSTVSGIASFLNPFGLAIGLLVIAGLVAAPFAFYRVVMOIRSNPM 881
Qy 782 KMIYPEIDKLKEREKSEIAPISE-----EELERIVLAMIHQONSHMETKTRK 830
Db 882 KALPYITTKALKNKAATSYGNEEDDGSDFDEAKLEAEAREMIKYMSMVSALEKQEKKAIK 941
Qy 831 DPKDS-----ILTRAQNMRLKRSQ--YSNLKNAESVE 860
Db 942 --KNSGVGLIASVSKLARRRGPKYTRLQONDME 975

```

RESULT 29

```

VGLB_HSVEA
ID_VGLB_HSVEA STANDARD; PRT; 980 AA.
AC P18551.
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B precursor (Glycoprotein 14).
GN GB OR GP14 OR 33.
OS Equine herpesvirus type 1 (strain ABI) (EHV-1).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
NCBI_TaxID=10328;

```

[1]

```

RN SEQUENCE FROM N.A.
RP Bonass W.A., Elton D.M., Stocks J.M., Killington R.A.,
RA Meredith D.M., Halliburton I.W.,
RL Submitted (JUL-1990) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M36298; AAA46068.1; --
DR InterPro; IPR000234; Glycoprot B.
DR Pfam; PF00606; Glycoprotein B; 1.
DR ProDom; PD000693; Glycoprot B; 1.
KW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 86
FT CHAIN 87 980
FT DOMAIN 87 852
FT TRANSMEM 853 870
FT DOMAIN 871 980
FT CARBOHYD 165 165
FT CARBOHYD 275 275
FT CARBOHYD 380 380
FT CARBOHYD 423 423
FT CARBOHYD 497 497
FT CARBOHYD 514 514
FT CARBOHYD 515 515
FT CARBOHYD 560 560
FT CARBOHYD 727 727
FT CARBOHYD 749 749
SQ SEQUENCE 980 AA; 109736 MW; 264D273CED70ESA1 CRC64;

```

Query Match 18.1%; Score 824; DB 1; Length 980;

Best Local Similarity 26.7%; Pred. No. 5.9e-43;

Matches 248; Conservative 161; Mismatches 385; Indels 134; Gaps 31;

```

Qy 32 VTPSPNTATWSTESPLTGHYTHSDSHGERGNNENRDSSEQNKNIYSPSPFPYRVC-SA 90
Db 83 VRAVPTTSPSPSTPTS--MSTH--SHG-----TVDPDLLPTETDPLRLAVRE 127
Qy 91 SGV-----GD-----VFRQTDHVCPDAS-DMVHSEGLILLIYKQNIIPFWRVK 134
Db 128 SGILAEADGDFYTCPPPTGTVVRIPPTCPKFDLGRNFTGIAVIFKENTAPIYKFRANV 187
Qy 135 YRKVVTSTVYNGIYSDSTNQHTFYKSIPEWETEK---MDTIYQCNS---LRLNTGG 187
Db 188 YKDIWTVRWKGYSHTSLSDR---YNDRPVSVSVEEIFGLIDSKGKSKAEYLR---D 240
Qy 188 NLLTYV---DRDDINMTVFLQGVDPDVKRYGSGQPELYLEPGWF-WGSYRRRTVNCE 243
Db 241 NIMHHAYHDEDEVELD--LVPSKFATPGARAWQTTNDTTSYVGMPWRHY-TSTVNCI 297
Qy 244 LMDMFARSNPFPFVATGDTVMSPFW-----SGEDDHENKMKHEKPMFVSVINNYKVD 299
Db 298 VEEVEARSVPYDSFALSTGDIVVASPFYGLRAAARIEHNSYAQDS---FRQVEGYRPRD 354
Qy 300 YQNRGTVPGLKTRIFLDREEVTLSEWKLKNMSYCPPLTLAKAFYNGIQTEHSGSYHFPAN 359
Db 355 LDSKLQAEPPVTKNFITTPHVTVSNWNTTEKKVEACTLTWKVEDELVRDEFRGYSYRTIR 414
Qy 360 DITASFTTSKEDMK-BENTTYHCLNEBEIKAEIEKKYAK-VNSTHSHKYGDLKYFKTDGGLY 417
Db 415 SISSTFISNTTQFKLESAPLTCVCVSKAEKAIDSIIYKKQYESTHVSFSDVEYILARGGFL 474
Qy 418 LVMQPLI-----QNRLDAKNKL-----NNETYSSRRRQAESTTDPMMEMT 459
Db 475 IAFRPMLSNELARLYLNLVRSNRTYDLKNLLPNANNNNTTTRRRSLLSVPEPQPTQD 534

```

```
QY 460 G-----NGAGGEYSSE-----NSITVAQVQYAYDNLIRINNI 492
Db |-----|-----|-----|-----|-----|-----|
QY 535 GVHREQLHLRKHRAVEATAGTSSNVTAQLELILKTTSSIEFAMQFAYDHIQSHVEM 594
Db |-----|-----|-----|-----|-----|-----|
QY 493 LEDLSKAWCREQRAALVWNELSKINPTSVMSIYNRPVSAKRIGDVISVNCIVVDQTS 552
Db |-----|-----|-----|-----|-----|-----|
QY 595 LSRIATAWCTLQNKERTLMNEVKINPSAIVSATLDERVAARVLGDVIAITHCAKI-EGN 653
QY 553 VSLHKSRLLSASDEKCFSPVPVTF---KEMNDSTIYKGGOLGVNNEILLTTTILETCQEN 609
Db |-----|-----|-----|-----|-----|-----|
QY 654 VYLQNSMR--SMDNCTCYSPPPVTFITTKNANRGSIEGOLGEBEIEFTERKULIEPCALN 711
QY 610 TETVFOAKTDMYIKNYEHLKTVPLSSITLDTFIALNFTLLENVDPKVIETLREKRL 669
Db |-----|-----|-----|-----|-----|-----|
QY 712 QKRYFKGKGVYVYENYTFVRKVPPTTEIVISYVELNLTLLEDREPLPLEVTRAELD 771
QY 670 SNVPDIETMPREYNYYAQRVSGRLKDLDLSTNRNQVDFAGSLMDDLGAVGQTVNNAVS 729
Db |-----|-----|-----|-----|-----|-----|
QY 772 TGLLDYSEIQRRNQLHALREFYDI--DSVNVNDNTAVIMOGIASFFKGLGKVGGEAVGLVL 829
QY 730 GVATLFSSIVTGFNFINKPFGGMLIIVVIGVLFAIYFLTKTIYETAPIKMIYPEID 789
Db |-----|-----|-----|-----|-----|-----|
QY 830 AAGAVVSTVSGIASFLNPFGLAIGLVIAAGFAFFAYRYVMQIRSNPMKALYPIIT 889
QY 790 KLERECKSEIAPISE-----EELERIVLAMHIHQONSHMETKTRKDPKDS--- 835
Db |-----|-----|-----|-----|-----|-----|
QY 836 -ILTRAQNLMLRKRSQ--YSNLKNAESVE 860
Db |-----|-----|-----|-----|-----|-----|
QY 948 LIASVNSKLALRRRGPKYTRLQNDTME 975
Db |-----|-----|-----|-----|-----|-----|

RESULT 30
VGLB HSAV1
ID VGLB HSAV1 STANDARD; PRT; 933 AA.
AC Q0463;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Glycoprotein B precursor.
GN Ul27.
OS Herpesvirus ateles (type 1 / strain Lennette).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae.
OX NCBI_TaxId=35243;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=93228440; PubMed=8385913;
RA Eberle R., Black D.;
RT "Sequence analysis of herpes simplex virus gB gene homologs of two
RT platyrrhine monkey alpha-herpesviruses.";
RL Arch. Virol. 129:167-182(1993).
CC -I- SUBUNIT: DIMER, PROBABLY LINKED BY DISULFIDE BONDS.
CC -I- SUBCELLULAR LOCATION: Integral membrane protein.
CC -I- MISCELLANEOUS: GB IS THE ONLY GLYCOPROTEIN THAT IS KNOWN TO BE
CC REQUIRED FOR VIRAL GROWTH.
CC -I- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC -I- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-30 IS THE INITIATOR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M95785; AAA43839.1; --
DR InterPro: IPR000234; Glycoprot_B.
DR Pfam: PF00606; Glycoprotein_B; 1.
DR ProDom: PD000693; Glycoprot_B; 1.
KW Glycoprotein; Transmembrane; Signal.
```

```
FT SIGNAL 1 31 POTENTIAL.
FT CHAIN 32 933 GLYCOPROTEIN B.
FT DOMAIN 32 733 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 754 769 POTENTIAL.
FT TRANSMEM 775 795 POTENTIAL.
FT TRANSMEM 798 818 POTENTIAL.
FT DOMAIN 819 933 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 107 107 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 161 161 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 418 418 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 450 450 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 697 697 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 747 747 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 933 AA; 101492 MW; 13A6CF88BB5A18DF CRC64;

Query Match 18.1%; Score 821; DB 1; Length 933;
Best Local Similarity 26.4%; Pred. No. 8; Se-43;
Matches 224; Conservative 164; Mismatches 378; Indels 84; Gaps 20;

QY 30 TGVTPPNTATWSTESPLTGHYTHDSSHGEBGN---ENRDSSEONKNIYSGSPSTFPY 85
Db |-----|-----|-----|-----|-----|-----|
QY 86 RYCSASGVGDVFRFQTDHVCPCD-ASDMVHSEGLLIYKQNIIPFMFRVKRYKVVTTSTV 144
Db |-----|-----|-----|-----|-----|-----|
QY 134 YVCPPTGATVQVFEPRPCDVAACKNFTGIAVVFKENIAPYKFTATKYKEITVSQT 193
QY 145 Y-----NGIYSOSITNQHTFYKSIAPWE--TEKMDTIYOCFNLSRLNTGNNLLTYVD 194
Db |-----|-----|-----|-----|-----|-----|
QY 194 WQGSRYLQTLGLYND-----RAPVPFSEITDLINGKGRCSRDVTVTRSORRTAVD 244
QY 195 RUDINMTVFLOPDVGTDPVKYSGQPELYLBPQFWGVSRYRRRTTVNCELMDMFARNPP 254
Db |-----|-----|-----|-----|-----|-----|
QY 245 GDEWGREVALVPAKTSPTNSRGWYTTDRVY-APNAHAGFYKGTGVVNCIVEEMEARSAFP 303
QY 255 FDFPVTATGDTVEMSPFWSGECDD--HENKMHKEKPEVSVINNY-KVYDVQNRCVTPLGKT 311
Db |-----|-----|-----|-----|-----|-----|
QY 304 YDSFVLATGEFVYVASFSGFSEDARRERNRYAPDRFQVDGVFPRLDQSGORATPV--V 361
QY 312 RIFLDREBYTSLWEKHLKNMSYCPYTLKAFYNGIOTHSYGVHVFANDITASFTTSKED 371
Db |-----|-----|-----|-----|-----|-----|
QY 362 RNLLTPTFTVGMDWKPRPNVCSVTKWQVVEEMVRAEYGSAPRFTSAALSATFTS--- 417
QY 372 MKEFNNTYH-----CLNEEIKAIEKKYA-KVNSTHXYKVDLKYKFTDGLGLVMO 421
Db |-----|-----|-----|-----|-----|-----|
QY 418 ----NLUTPPPELIEHSDCVAREAESIEAIYARRYNASHVRVGVQVYLAAGFFLAFQ 473
QY 422 PLIQNRLLDAKNLANNETYRSRSRROAESTTDPMMEMTNGAG--GEYSSENSITVAQVQ 479
Db |-----|-----|-----|-----|-----|-----|
QY 474 PLLSNSLAEMVRR---EALLGRSGDLAALAPPVVAAPASGAGPRCTISTTQTVFARLQ 530
QY 480 YAYDNLIRIRINILEDLSKAWCREQRAALVWNELSKINPTSVMSIYNRPVSAKRIGDV 539
Db |-----|-----|-----|-----|-----|-----|
QY 531 FTYDHIQKHVNEMLGRIAAAWQQLQOQLVNLWEARKLNPGAIASATVGTFRVARMGLDV 590
QY 540 ISVNCIVVDQTSVSLHKSRLLSASDEKCFSPVPVTFEKFMDNSTIYKGGOLGVNNEILL 599
Db |-----|-----|-----|-----|-----|-----|
QY 591 MAVSTCIPSPDNVIMQNSMR-IPGDPKTCYARPLVSFRYTDGELVEGOLGEDNEIRLE 649
QY 600 TTVLETQCBENTYFYFOAKTDMYIKNYEHLKTVPLSSITLDTFIALNFTLLENVDPKVI 659
Db |-----|-----|-----|-----|-----|-----|
QY 650 QNNVECTVGHKRYFVFGGVYVFEYAYSHQVSRAADVPVWSTFVDNLNLTMLDHEFLPL 709
QY 660 ELYTRDEKLSNVFDIETMPREYNYYAQRVSGRLKDLDLSTNRNQVDFAGSLMDDLGA 719
Db |-----|-----|-----|-----|-----|-----|
QY 710 EYVTRREIKSDGLDYAEVQRRNQLHALRFSIDIRIMND-SANA-ALMAGLARFFDGMGD 767
QY 720 VQOTVNVASGVATLFPSSIVTGFNFINKPFGGMLIIVVIGVLFAIYFLTKTIYETA 779
Db |-----|-----|-----|-----|-----|-----|
QY 768 AGAIGRAVLGVTEGLISVVGSSFLSNPFGALAVGLLVLAGLVAAAFAMRIMRLAN 827
QY 780 PIRMIYPEI-----DKLEREGKSEIAPISEELE-----RIVLAMHI 817
Db |-----|-----|-----|-----|-----|-----|
```

	Qy	105	CPDAS--DMVHSEGLILLYKONIIIPFMRVRKYRKVVTTSTVYNGIYSDSITNOHTFYKSI	167
	Dd	98	CPKFDLGRNFTEGIAVFKEKNIAFKPRANYYVDIVVKYWKGVSHTSLSDR---YNDR	154
	Qy	164	EPWETEKMDITY-----OCFNs---LRLNTGGNLLTYV---DRDDINMTVFLQPVDGVTPD	213
	Dd	155	VPSVSVEIFTLIDSKGKSSKAELYR----DNIMHHAYHDDEDEVELD--LVPSFEATPG	208
	Qy	214	VKRYGSQPELYLEPGWF-WGSYRRRTTNCCLMDMFARSNPFPDFVTATGDGTVMSPFW	272
	Dd	209	ARAWQTINDITTSYVGWMWRHY-IYSTVNCEIVEEARSVPYSDFSALSTGDIVTSPFY	267
	Qy	273	-----SGEDDHENKMHEKPWFVSVVINNVKVVDYQNRGTVPLCKTRIFLDREBYTLWSKHU	328
	Dd	268	GLRSAAQLHNYSQAQR---FRQVEGHQPROLDLSKLQAGEVTKGNFITTHPVTVSNMWE	324
	Qy	329	KNNISYCPLTLWAKAFYNGIGTESGSYHFVANDITASFTTSKEDMK-BENTTVHCUNEELK	387
	Dd	325	KXIEACTLTWKKEVDELVRDEFRGYSRPFTIRKSISTFLSINTQPKLEDAPLTDCVSKAEK	384
	Qy	388	AETEEKYAK-VNSTHSKYGDLKYFKPTDGGLVLVMOPLI-----QNRLLDAK	432
	Dd	385	DAIDSIVRKQVESTHVFGDVFFYLARGGFJLA FRPMISNELARLYLNELVRSNRTYDLK	444
	Qy	433	NKUN-NTY-SRRSRROAESTTD-----MMEMTCNGAGGESVSE-----	470
	Dd	445	NLLPNANHNTNRTRSLGISPETPQESLHRBQILHLRHKRAVEAANSTNSNNVTAQO	504
	Qy	471	-----NSTVAOVQVAYNLRIINNILEDLSKANCREOHRAALVWNLSKINPTSVM	523
	Dd	505	LELIKTTSSIEFAWLQFAYDHIOHSVNMJSRIAATACTLQNKRERTLMWVKVNPISAIV	564
	Qy	524	SMIYNRVPSAKRIDGVISVSNCI VVDOTSVSLHKSLRLLSASDEKCFGRPPVTF---	580
	Dd	565	SATLDERVARVLGDVIAITHCVKI -EGNVYLQNSMR---SDSNTCYSRPPVTFITKNA	621
	Qy	581	NDSTIYKGOLGVNNEILLTTTYLETQBENTYYPQAKTMVIYKXYEHKLTVPJSSITTL	640
	Dd	622	NSRGTIEGLEENEENVYTERKLI EPCAINKRFKFGKEYVYVENYTVVRKVPPTIEIVI	681
	Qy	641	DTPIALNFTILENVDPFKVELYTRDEXRLSNVFDIETMFREYNYAQORVSLGRKDLLDLS	700
	Dd	682	STVVELNLTLLEDDREFPLEVYTREALEDTGLLDYSEIQRRNQHALRFYDI--DSVNVN	739
	Qy	701	TNRNQFVDACFGLSMDDLGAQGTVVNAVSGVATLPSSITVTGFINFIKNPFPGMLMIIVI	760
	Dd	740	DNTAVIMOQTATPFKGLGKGEAVGTLVLGAGAVSTVSGIASFINNPFGLAICLLVI	799
	Qy	761	GVLPAIYFLTCKTKIYETAPIKMYPERI-----DKLKEREKGSEAPIASE-----EELE	809
	Dd	800	AGLVAFAFFAYRYVMQLRSNMPKALYPIPTRLSLKNKAKASYGONDODDTSDPDEAKLEBAR	859
	Qy	810	RIVLAMHIHQONSIMETKRDPK-D SILTRAQNMLRKRSg--YSNLKNAESVE	860
	Dd	860	EMIKYMSMVSALEKQEKKAMKKNKGVGLIASNVSKLARLRERGPKYTRLREDPDME	914
 RESULT 32 VGLB_HSVB2 STANDARD; PRT; 917 AA.				
	ID	VGLB_HSVB2	PRT;	917 AA.
	AD	P12641;		
	DT	01-OCT-1989 (Rel. 12, Created)		
	DT	01-AUG-1990 (Rel. 15, Last sequence update)		
	DT	16-OCT-2001 (Rel. 40, Last annotation update)		
	DE	Glycoprotein B-1 precursor.		
	OS	Bovine herpesvirus type 2 (strain BMV) (Bovine mammillitis virus).		
	OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;		
	CC	Alphaherpesvirinae; Simplexvirus.		
	OX	NCBI_TaxID=10296;		
	RN	[1]		
	RP	SEQUENCE FROM N.A.		
	RX	MEDLINE=88306231; PubMed=2841793;		
	RA	Hammerschmidt W., Conraths F., Mankertz J., Pauli G., Ludwig H.,		

KW Glycoprotein; Transmembrane; Signal.
 FT SIGNAL 1 86
 FT CHAIN 87 979
 FT DOMAIN 87 979
 FT TRANSMEM 852 869
 FT DOMAIN 870 979
 FT CARBOHYD 165 165
 FT CARBOHYD 275 275
 FT CARBOHYD 380 380
 FT CARBOHYD 423 423
 FT CARBOHYD 497 497
 FT CARBOHYD 514 514
 FT CARBOHYD 515 515
 FT CARBOHYD 560 560
 FT CARBOHYD 727 727
 FT CARBOHYD 749 749
 SQ SEQUENCE 979 AA; 110331 MW; 9A19866B791C5B36.CRC64;

Query Match 17.1%; Score 777.5; DB 1; Length 979;
 Best Local Similarity 26.3%; Pred. No. 4.2e-40;
 Matches 247; Conservative 159; Mismatches 377; Indels 155; Gaps 34;

QY 32 VTPSPNTATWSTESPLTGHYTHDSHGCGNNEIRDSEEQKNLYGSPFPYRVC-SA 90
 DB 83 VRAVETSPPTPTS--MSTH--SHG-----TVDPDLLPTETDPLSLAVRE 127
 QY 91 SGV-----GD-----VRFOTDHVCPDAS--DMVHSEGLIILYKONIIPFMFRVK 134
 DB 128 SGILAEADGDFYTCPPPTGTVVRIEPPRTCPKFDLGRNFTGIAVIFKENTAPYFRANV 187
 QY 135 YRKVTTSTVYNGIYSDSITNHTYKSIETPWEK-----MDTIYQCENS---LRLNTGG 187
 DB 188 YKDIWTVRWKGYSHTSLSDR---YNDRPVPSVEEITGLIDSKGCKSKAEYLA---D 240
 QY 188 NLLTV---DRDDINMTVLPQDVGTVDPVKRYGQPELYLEPGWF--NGSVRRRTTVNCE 243
 DB 241 NIMHAYHDDDEVELD--LVESKFPATPGARAWQTTNDTTSYVGMWPRHY--TSTSVNCI 297
 QY 244 LMDMEARSNPPDFVTATGDTVEMSPFW---SGEDDHENKMKHKKPWFVSVINNYKVD 299
 DB 298 VEEVEARSVPYDVSALSTGDIYVSPYGLRAAARIENHNSVAQER---FRQVEGYRPRD 354
 QY 300 YQNRGTVPGLKTRIFLDREEVYLSWEKHLKNMSYCPJTLWKAFYNGIQTHSGSYHFVAN 359
 DB 355 LDSKLQAEPEPTKNEFITPHVTVSNWTEKKVEACTLTWKVEVDDELVRDEFRGSYRFTIR 414
 QY 360 DITASFTTSKEDMK--EFNTHYHCLNEEIKAEIKKYAK--VNSTHSGYGLKYFTDGLY 417
 DB 415 SISSYFISNTTQFKLESAPLTCVSKAEKAIDSYKQYESTHVFSGDVEYLLARGGFL 474
 QY 418 LWQPLI-----QNRLLDAKNKL-----NNETYSRRSRQAESTTDPMEWT 459
 DB 475 IAFRMLSNELARLYNLVLSNRNYDYLKLLNPNANNNNNTTTRRRSLLSVPEPQPTQD 534
 QY 460 G-----NGAGGEYSSE-----NSITVAQVQYAYDNLRINNI 492
 DB 535 GVHREQLHLRLHKRAVEATAGTDSNVVAKQLELIKTTSSIEFAMQLQAYDHIQSHVNM 594
 QY 493 LEDLSKANCREQHRAALVWNLKSNKPNVSMYVNRVPSAKRIGDVTSVNCIVDQTS 552
 DB 595 LSRIATAWCPQNKERPLWNEVMKTIPTSAIVSATLDERVAARVLGDVIAITHCAKI--EGN 653
 QY 553 VSLKSLRLLASDEKCFSPDPVTF---KFNWDSIIYKGLGVNNEILLTLYLETQEN 609
 DB 654 VYLQNSMR--SMDSTCYSRPPTVTITTKNANNRSGIEGQGEENEIEFTERKLIEPCALN 711
 QY 610 TEYFPAKTDWYIYKNYHLKTVPLSSITTLDTFIALNFTLLENVDKVIELYTRDEKRL 669
 DB 712 QKRYFKCKEYVYENYTFVRKVPTEIEVISTYVELNLTLLEDEFLUPLLEVYTRAELED 771
 QY 670 SNVFDIETMFREYNYAQRVGLRKDLLDLSTNRNQFVD-----AFGSLMDLIGA 719
 DB 772 TGLLDYSIORRNQHLALR-----FYDIDSVVN--VDNTAVIIRGSPAFSAWVKWR 822

QY 720 VQOTVNAVSGVATLFFSSIVTGFINFKNPFGLMLIIVIGVLFAIYFLTKTKIYETA 779
 DB 823 PWERSFSARGAVV-----STVSGIACFLNNPFGGLIAGLVAAFAFYVYVQIRSN 878
 QY 780 PIKMIYPEIDKLKREKSEIAPISE-----EELERIVLAMIHQONSHMETKT 828
 DB 879 PMKALYPITTKALKNKAITSYQNEEDDGSDFDEAKLEAREMIKYMVMSVALEKQEKKA 938
 QY 829 RKDPKDS---ILFPAQNMVKRSG---YSNLKNAESVE 860
 DB 939 IK--KNSGVGLIASNVSKLALRRRGPKYTRLQONDIME 974

RESULT 34
 VGLB PRIVIF
 ID VGLB PRIVIF STANDARD; PRT; 913 AA.
 AC P08355;
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glycoprotein GII precursor.
 OS Pseudorabies virus (strain Indiana-Funkhauser / Becker) (PRV).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.
 OX NCBI_TaxID=31523;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87284141; PubMed=3039163;
 RA Robbins A.K., Dorney D.J., Wathen M.W., Whealy M.E., Gold C.,
 RA Watson R.J., Holland L.E., Weed S.D., Levine M., Glorioso J.C.,
 RA Enquist L.W.;
 RT "The pseudorabies virus gII gene is closely related to the gB
 glycoprotein gene of herpes simplex virus.";
 RL J. Virol. 61:2691-2701 (1987).
 RN [2]
 RP SEQUENCE OF 847-913 FROM N.A.
 RX MEDLINE=89279298; PubMed=2543777;
 RA Simon A., Mettenleiter T.C., Rziha H.J.;
 RT "Pseudorabies virus displays variable numbers of a repeat unit
 adjacent to the 3' end of the glycoprotein gII gene.";
 RL J. Gen. Virol. 70:1239-1246 (1989).
 RN [3]
 RP EXPORT PATHWAY.
 RX MEDLINE=90219190; PubMed=2157862;
 RA Whealy M.E., Robbins A.K., Enquist L.W.;
 RT "The export pathway of the pseudorabies virus gB homolog gII involves
 oligomer formation in the endoplasmic reticulum and protease
 processing in the Golgi apparatus.";
 RL J. Virol. 64:1946-1955 (1990).
 CC - SUBUNIT: DIMER, PROBABLY LINKED BY DISULFIDE BONDS.
 CC - SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb.ch/announce/>
 or send an email to license@sib-sib.ch).
 DR EMBL; M17321; AAA47465.1; --
 DR EMBL; D00464; BAA00359.1; --
 DR PIR; A29159; VGBEPS.
 DR InterPro; IPR000234; Glycoprot B.
 DR Pfam; PF00606; Glycoprotein B; 1.
 DR ProDom; PD000693; Glycoprot_B; 1.
 KW Glycoprotein; Transmembrane; Signal.
 FT SIGNAL 1 40
 FT CHAIN 41 913
 FT DOMAIN 41 750
 FT TRANSMEM 751 819
 FT DOMAIN 820 913
 FT GLYCOPROTEIN GII.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).

```
FT CARBOHYD 151 151 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 261 261 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 516 516 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 573 573 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 633 633 N-LINKED (GLCNAC. . .) (POTENTIAL) .
FT CARBOHYD 697 697 N-LINKED (GLCNAC. . .) (POTENTIAL) .
SQ SEQUENCE 913 AA; 100234 MW; 5D560D235E956437 CRC64;

Query Match 17.1%; Score 775.5; DB 1; Length 913;
Best Local Similarity 27.0%; Pred. No. 5.1e-40;
Matches 235; Conservative 165; Mismatches 388; Indels 81; Gaps 26;

Qy 26 SIATGTVSPNTATWSTESPLTCHYTHDSHGCGNNENRDSBEQKNLYGSPSPFPY 85
Db 85 SLEIEAFSP-----GSEAP-DGEYGLDARTAVRAAATERDR-----F 123

Qy 86 RVCSAGSGVDVFRPQTQDHVCPDASD-MVHSEGIILYKQNIIPFMFRVKYKRVVTTSTV 144
Db 124 YVCPPPSGSTVVRLEPEQACEPYSGQRNFTGIAVLKFNIAHPKFAHIYKXNVITTV 183

Qy 145 YNGIYSISITNQHTYKSIKBPWE-TEKMDTYYQCFNSLRNLNTGNNLLTYVDRDDINMTVF 203
Db 184 WSGSTYAAITNRFTDRVPVPVQEITDVIDRRGCKVSKAEYVRNNHKVTAFFDRDENPVEVD 243

Qy 204 LQPDVGVTPDKVYSGOPELVLEPGWFGSYRRRTVNCCLMDMFABSNPPDFEFTVATG 263
Db 244 LRPRNLALGTRGWHHTNDYTKIG-AAGFVHTGTSVNCIVEEARSVPYDYSFALSTG 302

Qy 264 DTVMSPFWS-GEDDHENKMEKFPWFSVINNYKVQDVYQNRGTVPGLKTRIFIDREBYTL 322
Db 303 DIVTMSFPYGLREGAGHEGICYAQRFOQVEHYYPIDLDLSPLRASESVTRNFLTPTFTV 362

Qy 323 SWEKHLNMSYCPITLWKAFFNGIQTE-HSGSYHFVANDITASPTT--SKEDMKFNTTY 379
Db 363 AWDWAPKTRRYCSLAKWEAREEMTRDETROGSFRFTSRALGASVDSVDTQLDRVHLG- 421

Qy 380 HCLNEEIKAEIEKKY-AKVNSTHSGYGL-KYFKTDGGLVLMQPLIQNLRLDA-KNKLN 436
Db 422 DCVUREASEALDAILYRRYNSHTVLADRPVEYLARGGVVAFRPLISNELAQYARELE 481

Qy 437 N-----ETYSRRSRQ--AESTDTPMMEMTNGAGGEYSSENSITVAQOVAYD 483
Db 482 RLGLAGVGPAPAAARRARRSPGACTPEP---PAVNGTGLRLITTSAEFARLQTYD 538

Qy 484 NLRIRINILDLKAWCRQHRALVWNLSEKINPTSMYIYVRRPVSAKRIGDVISVS 543
Db 539 HIQAVNDMLGRIAAAACELQNKDRTLWSEMSRLNPSAVATAALGQVSARMGLGDVMAIS 598

Qy 544 NCIVVDQTSVSLKSLRLSSADEKCFSPRPVTPKFMNDSTIYKGLGVNNEILLTLYL 603
Db 599 RCVEV-RGGVTVQNSMR-VGERGTCYSRPLVTEH--NGTVIEGGQGDNDNELLISRDLI 655

Qy 604 ETCOENTYFQAOKTDMYIYKNYEHLKTVPLSSITLDTFIALNFTLLENVDKVIELYT 663
Db 656 EPTGNGHRRYFKLGSYVYEDYNYVRMVEPE--TISTRTVNLTLLEDEFLPLEVYT 713

Qy 664 RDEKRLSNVFDIETMFREYNYA-----QRVSLGRKDLDDLSTNRNQFVDAGSLMDDL 717
Db 714 REELADTGLLDYSEIQRRNLHAKFYDIDRVVKVDHNVLLRGIANFF-----QGL 765

Qy 718 GAVGQTVNNAVSGVATLFSISITGTFINFIKNPFGMLMIIVIGVLPAIFYLTKTKIYE 777
Db 766 GDVGAAGVKVVLGATGAVIYAVGGMVGSFLNPGALGILVLVLGVLAAFLAYRHISRLR 825

Qy 778 TAPIKMTYPIDLKEREKSEIAPISEEBLER-----IVLAMHIHQONSHMETKT 828
Db 826 RNPMAKALYPVTTKLKEDGVD-CDVDEAKLDQARDMIRYMSIVSAL---EQEHRKARK 881

Qy 829 RKPDKUSILTRAQNMILKRSGYSLNKNAB 857
Db 882 NSGPA-LLASRVGAMATRRRHHYQRLSESD 909
```

RESULT 35

VGLB_VSVBP STANDARD; PRT; 928 AA.

AC P17471; 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein I precursor (Glycoprotein GVP-6) (Glycoprotein 11A)
DE (Glycoprotein 16) (Glycoprotein G130).
GN GI.

OS Bovine herpesvirus type 1 (strain P8-2).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10324;
RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=89020821; PubMed=2845660;

RA Misra V., Nelson R., Smith M.

RT "Sequence of a bovine herpesvirus type-1 glycoprotein gene that is

homologous to the herpes simplex gene for the glycoprotein gB.";

RL Virology 166:542-549(1988).

CC -!- SUBUNIT: DIMER, PROBABLY LINKED BY DISULFIDE BONDS.

CC -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC -----

DR EMBL; M23257; AAA46013.1; --

DR PIR; A31166; VGBEBG.

DR InterPro; IPR000234; Glycoprot B.

DR Pfam; PF00606; Glycoprotein B; 1.

DR ProDom; PD000693; Glycoprot_B; 1.

KW Glycoprotein; transmembrane; Signal.

FT SIGNAL 1 67

FT CHAIN 68 928 GLYCOPROTEIN I.

FT DOMAIN 68 780 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 781 801 POTENTIAL.

FT TRANSMEM 804 824 POTENTIAL.

FT DOMAIN 825 928 CYTOPLASMIC (POTENTIAL).

FT CARBOHYD 105 105 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 484 484 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 579 579 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 637 637 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 703 703 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CONFLICT 409 409 S -> T (IN REF. 1; AAA46013).

FT CONFLICT 673 673 G -> P (IN REF. 1; AAA46013).

SQ SEQUENCE 928 AA; 102177 MW; B47982224FCD769D CRC64;

Query Match 15.9%; Score 722; DB 1; Length 928;

Best Local Similarity 26.2%; Pred. No. 1e-36;

Matches 240; Conservative 156; Mismatches 427; Indels 94; Gaps 26;

Qy 2 AGSLKRGSVLALWLYQVALYSLSIAETCVT-----SPPNTATWSTESPLTGHY 51

Db 41 AGGARAALAAALLWATWALLAAAPAGRPATTPAPPPEEAASPAASPSPGPGDDA 100

Qy 52 GTHDSSHGGRGNNENRDSBEQKNLYGSPSPFPYRVCSAGSGVDVFRFQTDHVCPO-ASD 110

Db 101 ASPDNSTDVRAALRLQAAGENSRRF-----VCPSPSGATVVLAPARCPCEVGLG 151

Qy 111 MVHSEGIILYKQNIIPFMFRVKYKRVVTTSTVYNGIYSDSITNOHTYKSTPEWE-TE 169

Db 152 RNYTEGIVYKENIAPYTFKAIYYKXNVITTWAGSTYAAITNQYTRVPVGMGEITD 211

Qy 170 KMDTIYQCFNSLRNLNTGNNLLTYVDRDDINMTVFLQPDVGVTPDKVYKSGOPELVLEPGW 229

```

Db 212 LVDDKWKRLSKAABYLRSGKVVAFDRDDDPWBAPLKPARLSAPGVRGWHHTDDVYALG- 270
Qy 230 FWGSYRRRTTVCNCLMDMFARPNPPDFPVATGDTVMSPWS--GEDDHENKMEKEWPF 288
Db 271 SAGLYRTGTSVNCIVEEYVARSVPYDPSALSTGDIYMSPPYIGREGAHRHTSYSPER 330
Qy 289 VSVINNYKVVDYQNGRTVPLGKTRIFLDRREYTLSEWHLKNNMSYCPILTMKAFYNGIQ 348
Db 331 FQIEGYKRDMMATGRRLKEPVRNFRFLRQHTVAVMDVPRKNCVSLAKWREADMLRD 390
Qy 349 EHSYSYHVFANDITASFTTSKEDMKEFNNTY--HCLNEELKABIEKYY--AKVNSTHSGVD 406
Db 391 ESRGNFRFRARSLSATFVSDSHTFALQNVPLSDCVIEEAAEAVERVRYNGTHTVLGSG 450
Qy 407 LKYFKTDGGLYLWVMOPLQNRLLDAKNKL-----NNETYS-----RRRRQAE 449
Db 451 LETYLARGGFVAFPRMLSNEL--AKLYLOELARNGTLEGLFAAAAPKPGRRARRPRR 508
Qy 450 STTDPMMEMTNG-----AGBYSSENSITVAQVQYAYDNLRIINNILEDLSKAWC 501
Db 509 LRPAP-----GRQARRRRHAGGRVTVSLAEFAALQFTHDTRTSEHHV--HRLASPM 562
Qy 502 RQHRALVWNLKINPTSMVSMIYNRPVSARKIG-----DVI--SVNCIVVDQTSVSL 555
Db 563 LLQNKERALWAEAKLNPSAAAASAALDRPPRACWGTHGRDVLPRAGRGALFIENSARA 622
Qy 556 HKSLRLLSASDEKCFSRPPVTFKFNDSTIYKQGLGVNNEILLTTTLYLETCOENTYFQ 615
Db 623 RR--RL-----QPRFLSFQNESEPVGQGEDNELLPGRELVEPCPTANKHRYFR 671
Qy 616 AKTDMYIYKNYBHLKTVPLSSITTLDTFIALNFTLLENVDFKVIELYTRDKRLSNVPDI 675
Db 672 FGADVVYVYENYVRRVPLAELEVISLTFVGLNLTVLEDBREFLPLEVYTRAEALADTGLDY 731
Qy 676 ETMFREYNYAQRVSGRLKDLDLSTNRNQFVDAFGSLMDLGLAGVQGVTVNNAVSQATLF 735
Db 732 SBIQRNQLHELRYFYDI--DRYVKTGDGNAIMRGLANFPQGLGAVGQAVGTVVLGAAGA 789
Qy 736 SSIIVGTFINFKPFGGMLIIVGLVFAIYFLTKTKIYETAPIKMIYPEIDK--LKER 794
Db 790 LSTVSGIASFIANPFGALATGLLVLAAGLVAALAYRYSRUSRNPMKALYPIITRALKDD 849
Qy 795 ECKSEIAPISEE-----LER---IVLAWHIHQONSHMETKTRDKPKDS--ILTR-AQN 842
Db 850 PCNRN--PGEEREEFDAKLEQAREWIKYMSLVSAVERQEHKAKSKAARLLATRLTQL 907
Qy 843 MLRKS--GYSNLKNAE 857
Db 908 ALRRRAPPEYQQLPMAD 924

```

RESULT 36 RBP2_PLAVB

```

ID RBP2_PLAVB STANDARD; PRT; 1251 AA.
AC Q00799;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Reticulocyte binding protein 2 (Fragment).
GN RBP2.
OS Plasmodium vivax (strain Belem).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=31273;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92315338; PubMed=1617731;
RA Galinski M.R., Medina C.C., Ingravallo P., Barnwell J.W.;
RT "A reticulocyte-binding protein complex of Plasmodium vivax
RT merozoites."
RL Cell 69:1213-1226(1992).
CC -!- FUNCTION: INVOLVED IN RETICULOCYTE ADHESION. SPECIFICALLY BINDS TO
HUMAN RETICULOCYTE CELLS.

```

```

CC -!- SUBCELLULAR LOCATION: Membrane-bound (Probable).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; M88098; AAA29744.1; -.
KW Malaria; Receptor; Membrane.
FT NON TER 1
FT NON TER 1251
SQ SEQUENCE 1251 AA; 143741 MW; 54BA51C7404AC572 CRC64;

Query Match 3.2%; Score 144; DB 1; Length 1251;
Best Local Similarity 19.0%; Pred. No. 0.46;
Matches 165; Conservative 138; Mismatches 319; Indels 246; Gaps 42;

Qy 109 SDMVHSGILLIYKONI-IPMPRV-----RKVRKVVY---TSTV--YNGIYSDSITN 155
Db 506 STIINIGALKESKGNVEIGFLEKLEIGKRNKRLKVDITKKSINSTVGNFSSLFNNFDLN 565
Qy 156 QHTFYKSIETPMEKMDTIYQCF-----NSLRNTGGNLLTYVDRDDIN 199
Db 566 QYDENKNIYDYE-NKMGIEYNEFGSLNKISENLNASENTSDYNSAKTLRLAEQKEKVN 624
Qy 200 MTVLEQVQVGPVPRKYG-----QPELYLEPGWFGWSYRRRTTV 240
Db 625 LLNKEEEANKYLRDVKKVESFRPIFNKESLDKINEMIKKQLTVNEG--HGNVKQLVEN 682
Qy 241 NCELMDMFARPNPPDFPVATGDTVMSPWSEDDHEN--KMEKPMFWFVSVINNYKVD 299
Db 683 IKELVD-----ENLUSDILKQATGKNEIQKI-----THSTLNKAKTILGHVDTSKATVG 733
Qy 300 YQNRG-----TVPLGKTRI-----FLDREBYTLSEWHLKNNMSYCPILTMKAFYNGIQ 348
Db 734 IKITPELALTELLGDAKLKTAQELKFESKNNVLETEMNKNTELDV-----HKNIQD 787
Qy 349 EHSYSYHVFANDITASFTTSKEDMKEF-----NTTY---HCLNE-----EIKABIEKYYA 395
Db 788 AYKVALBILAH--SDEIDTKQKSSKLIEMGNQYILKVVLINQYKNKISSIKSEEAHSV 845
Qy 396 KVNSTHSGYGLKYF-----KTDGGLYLVMOPLQ-----NRLLDKNNKNNETYSRRSRQ 447
Db 846 KIGNVSKHSELSEKITCSDKSYNIIALEKQTELQNRNFTQETKNTNSDKLEKIKTD 905
Qy 448 AESTTDPMMEMTNGAGGEYSSENSITVAQVQYAYDNLRIINNILEDLSKAWCREQHRA 507
Db 906 FESLKNALKTLEGEVNAKASDN-----HEH-- 932
Qy 508 ALVWNLKSNINPTSMVSMIYNRPVSARKIGDVISVNCIVVDQTSV--SLHKSLRLLSASD 566
Db 933 --VQSKSEPVNP-----ALSE--IEKEETDIDSLNTALDELKKG 968
Qy 567 EKC-FSRPPVTFKEMNDSTIYKQGLGVNNEILLTTTLYLETCQENTYVYFOAKTDMYIKN 625
Db 969 RTCEVSR---YKLKIDT-----VTKEISDDTELINTIEKNVAYL-----AYIKKN. 1011
Qy 626 Y-----EHLKTVPLS-----SITLDTFIALNFTLLENVDFKVIELY 662
Db 1012 YEDTVQDVLINHEFNTKQVSNHEPTNFDKSNKSSEBELTKAVTDSKTIISKLGKVIIEVN 1071
Qy 663 TRDEKRL--SNVFDIETMFREYNYAQRVSGRLK-----DLDDLTNRNQFVDA--- 709
Db 1072 ENTEMNTIESSAKEIEAL---YNELKNKTSLSNEIYQTSNEVKLQEMKSNADKVIDVSKI 1128
Qy 710 FGSLMDLGLAGVQGVTVNNAVSQVATLTFSSIVTG--FINFKNPPFGGMLIIVGLVPAIY 767
Db 1129 FNTVLD-----TQKSNIVTNOHNSNNVKNKDKLQELQELIDSSFTLE 1171
Qy 768 FLTKTKIYETAPIKMIYPEIDKLEREGKSEIAPISEEELERIVLAWHIHQONSHMETK 827

```

```
Db 1172 SIKKFEIY--SHIKTNGIEQL-QQTNKSEHNVAKHK-EKIV---HLINRVESLKG 1224
Qy 828 TRKDPKDSILTRAQMLRKSGVSNLKN 855
Db 1225 VKNHDDDOYMKKLNASLN---DNIKN 1248

RESULT 37
KIPI_YEAST
ID_KIPI_YEAST STANDARD; PRT; 1111 AA.
AC P28742;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Kinesin-like protein KIPI.
GN KIPI OR CIN9 OR YBL063W OR YBL0504 OR YBL0521.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=92317166; PubMed=1618910;
RA Roof D.M., Meluh P.B., Rose M.D.;
RT "Kinesin-related proteins required for assembly of the mitotic
RT spindle.";
RL J. Cell Biol. 118:95-108(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=94205266; PubMed=8154187;
RA Scherens B., el Bakoury M., Vierendeels F., Dubois E., Messenguy F.;
RT "Sequencing and functional analysis of a 32,560 bp segment on the
RT left arm of yeast chromosome II. Identification of 26 open reading
RT frames, including the KIPI and SEC17 genes.";
RL Yeast 9:1355-1371(1993).
RN [3]
RP CHARACTERIZATION.
RC STRAIN=S288C;
RX MEDLINE=92354062; PubMed=1643659;
RA Saunders W.S., Hoyt M.A.;
RT "Kinesin-related proteins required for structural integrity of the
RT mitotic spindle.";
RL Cell 70:451-458(1992).
CC -!- FUNCTION: REQUIRED FOR ASSEMBLY OF THE MITOTIC SPINDLE. INTERACT
CC WITH SPINDLE MICROTUBULES TO PRODUCE AN OUTWARDLY DIRECTED
CC FORCE ACTING UPON THE POLES. FOLLOWING SPINDLE ASSEMBLY, CIN8 AND
CC KIPI APPARENTLY ACT TO OPPOSE A FORCE THAT DRAWS SEPARATED POLES
CC BACK TOGETHER. THIS FORCE SEEMS TO BE MEDIATED BY KAR3.
CC -!- SUBUNIT: MIGHT BE DIMERIC.
CC -!- SUBCELLULAR LOCATION: SPINDLE MICROTUBULES THAT LIE BETWEEN THE
CC POLES.
CC -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. BIMC
CC SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
DR GO; GO:0007020; P:microtubule nucleation; IPI.
DR GO; GO:000092; P:mitotic anaphase B; IMP.
DR InterPro; IPR001752; kinesin_motor.
DR Pfam; PF00225; kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
DR PROSITE; PS00667; KINESIN MOTOR DOMAIN2; 1.
KW Motor protein; Cell division; Microtubules; ATP-binding; Coiled coil;
KW Mitosis; Cell cycle.
FT DOMAIN 50 423 KINESIN-MOTOR (BY SIMILARITY).
FT DOMAIN 424 510 COILED COIL (POTENTIAL).
FT DOMAIN 648 670 COILED COIL (POTENTIAL).
FT DOMAIN 710 780 COILED COIL (POTENTIAL).
FT DOMAIN 808 828 COILED COIL (POTENTIAL).
FT NP_BIND 141 148 ATP (BY SIMILARITY).
SQ SEQUENCE 1111 AA; 125794 MW; 212F8279766137FC CRC64;

Query Match 3.1%; Score 140; DB 1; Length 1111;
Best Local Similarity 21.1%; Pred. No. 0.69;
Matches 121; Conservative 98; Mismatches 199; Indels 156; Gaps 29;

Qy 232 GSYRRTTNCCLMDMFARSNPPDFVTATGDTVMSPPWSGEDDHENKMKHKPMFVSV 291
Db 264 GSLKRKVAAT-KCNDLSSRSHTVF-----TITTNIVEQ-----DSKDHQKNFVK----- 308
Qy 292 INNYKVVDYQNRGTVPGLKTRIFLDREEVYLSWEKHLKMSYCPILTWKAFYNGIQTEHS 351
Db 309 IGKLNVLVDL-----AGSENINRSGAENKRAQEAGLINKSL--LTGLRVI--NAL--VDHS 357
Qy 352 GSYHFVANDIT-----ASFTTSKEDMKFNTTYHCLNEIEKAEIKYA 395
Db 358 NIPIPRESKLTLLQDSLGMYKTCIIATISPAKISMEETAST---LEVATRAKSIKNTP 414
Qy 396 KVNSTHSKYGDLYKFTDGLYLVWQPLIQNRLLDAKNLNNTSYRRRRRQAEISTDPM 455
Db 415 QVNQSLSKDTCLKDY-----IQBIEKLRNDLKNSRNKQGFITODQL 456
Qy 456 MMTGNGAGGEYSSENSITVAQVOYAYDNLR-----IRINILEDLSKAWCR 502
Db 457 -----DLYESNILIDEQNLKHNLRQIKFKENYLNQLDINNLLQS-----E 500
Qy 503 EQHRAALVWN-----ELSKINPTSVMSG---IYNRPVSAKRIGDIVSVSNICIVVD 549
Db 501 KEKLAIQNFVDFSNFYSEIQKHHTNLELMNEVIQORDFSLENSQKYNTNQNMQL- 559
Qy 550 QTSVSLHKSRLRLSAS-----DEKC-----FSRPPVTFKFMNDSTIYKQQLGVNNE 595
Db 560 KISQVQLTTLTQLGSLNNYNSKCSVIKGVTEELTRNVNTHKAKHDSL-KSLNITTN 618
Qy 596 ILLT-----TTYLETQ-ENTYFYQAQTDY-----IYKNYEH-----LKTVPIS 635
Db 619 LLNMNMNEILVRISLTSLEIFQSDSTSHYRKDLNEIYQSHQQFLKNIQNIDIKSCLDISGS 678
Qy 636 SITTLDTFIALNFT-----LLENV---DFKVIETYTRDEKRLSNVFDIETMFREYN 684
Db 679 ILTSINE-ISOCTTNLNSMNVLIENQSQSGSKLKEQDLKIKLKN--DLINERRISNQ 735
Qy 685 YAQRVSGLRKDLDD-LSTNRNQFVDAFGSLMDDL 717
Db 736 FNOQLAEMKRYQDHSVSRTRSEFHDNLKNCIDNL 769

RESULT 38
CUT7_SCHPO
ID CUT7_SCHPO STANDARD; PRT; 1085 AA.
AC P24339;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Kinesin-like protein cut7.
GN CUT7 OR SPAC25G10.07C.
OS Schizosaccharomyces pombe (Fission yeast).
```


OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE=91015362; PubMed=2145514;
 RA Hagan I., Yanagida M.;
 RT "Novel potential mitotic motor protein encoded by the fission yeast
 RT cut7+ gene.";
 RL Nature 347:563-566(1990).
 RN [2]

RP SEQUENCE FROM N.A.
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Guillermo R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouras J., Peat N., Hayles J., Baker S., Basmah D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy I., Niblett D., Odell C.,
 RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Grynoprez B.,
 RA Weljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Leinach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gallard C., Tallada V.A., Carlson A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomycetes pombe";
 RL Nature 415:871-880(2002).

CC -! FUNCTION: COULD BE A SPINDLE POLE BODY MOTOR. ON TRANSITION FROM
 CC G2 TO M PHASE OF THE CELL CYCLE, THE SPINDLE POLE BODY DUPLICATES;
 CC THE DAUGHTER POLE BODIES SEED MICROTUBULES WHICH INTERDIGITATE TO
 CC FORM A SHORT SPINDLE THAT ELONGATES TOWARD THE NUCLEUS AT
 CC METAPHASE. MUTATIONS AT CUT7 BLOCK SPINDLE FORMATION.
 CC -! SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. BIMC
 CC SUBFAMILY.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; X57513; CAA04738.1; --
 CC EMBL; Z70691; CAA94636.1; --
 CC F01; T38378; T38378.
 CC HSP; P17119; 3KAR.
 CC GeneDB Spombe; SPAC25G10.07c; --
 CC InterPro; IP001752; kinesin_motor.
 CC Pfam; PF00225; kinesin; 1.
 CC PRINTS; P00360; KINESINHEAVY.
 CC SMART; SM00129; KISC; 1.
 CC PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
 CC PROSITE; PS00412; KINESIN MOTOR DOMAIN2; 1.
 CC Motor protein; Cell division; Microtubules; ATP-binding; Coiled coil;
 KW Mitosis; Cell cycle; Phosphorylation; Repeat.
 FT DOMAIN 70 435 KINESIN-MOTOR (BY SIMILARITY).
 FT COILED COIL (POTENTIAL).
 FT COILED COIL (POTENTIAL).
 FT DOMAIN 436 604
 FT DOMAIN 715 740

FT DOMAIN 897 955
 FT NP_BIND 159 166
 FT REPEAT 987 998
 FT REPEAT 999 1010
 FT MOD_RES 1011 1011
 FT CONFLICT 34 61
 FT PHOSPHORYLATION (BY CDC2) (BY
 FT SIMILARITY).
 FT SASNPKRRRPPPTDTGYPDRSDTNSPT ->
 FT LRAILGNDVSLLLTL (IN REF. 1).
 SQ SEQUENCE 1085 AA; 122133 MW; 5669277875559D58 CRC64;

Query Match 3.1%; Score 139; DB 1; Length 1085;
 Best Local Similarity 19.9%; Pred.No. 0.77;
 Matches 164; Conservative 106; Mismatches 311; Indels 244; Gaps 41;

QY 152 SITNQHFYKIEPWETEKMDTIYQCFNSLRNLNTGNTLLTVVDRDDINMTVFLQPDGVT 211
 DB 26 STPNSH-FRSASNPKRRPPPTI-----DTG-----YDSDSTN-----SPTDHAL 65
 QY 212 PDVKRYGSGQPELYLEPGFMGYSRRRTTVNCEL-----MDMPARSN----- 252
 DB 66 HD-----ENEININVVVRGRTDOEVRDNSSSLAVSTSGAMGAE 104
 QY 253 -----PPDPFVTATGDTVEMSPFSGEDDH-----ENKMHEKPFVSVINNYK--VVDYQ 301
 DB 105 LAIQSDPSSMLVT---KTYAFDKVFGPEADQMLFENSV--APMLEQVLNGYINCTIFAYG 159
 QY 302 NRGTVPLGKTRIFLDREYETLSWEKHLKNMSYCPILTWKAFYNGIQTSHSGSYHFVANDI 361
 DB 160 QTGT---GKT-----YTMGSD-----LSDSDGILSEGAGLIPRALYQL 194
 QY 362 TASPTTKSKEDMKFNTTVHCLNEEIK---AEIEKKYAKVNSTHSGYDGL-----KYF 410
 DB 195 FSSLDNSNQEVAVKSYVELNNEIRDLVSEELKPARVFEITSRRGNVITGIEEYI 254
 QY 411 KTDG-GLYLVWQPLIQNRLLDAK-NKLNNEYTS-----RRSRQEAESTTDPMMWMTGNAG 464
 DB 255 KNAGDGLRLREGSHRRQVAATKCNLSSRSHTFTITLHRKVSGMTDETNSLTINNS 314
 QY 465 GEYSSENSITVAQV-----QYAYDNLRIK-----INNILEDLSK---AWCREQHRALV 510
 DB 315 DDLRASKLHMDVLGASENIGRGAENKRAETGMINQSLTLGRVINALVEKAHPIYR 374
 QY 511 WNELSKINPTSVMSIMYNPVSAKRIGDVISVSNCI VVDQTSVSLHLSRLLS-ASDEKC 569
 DB 375 ESKLTRLQDS-----LGKTKTSMIVTSSNTNLEETISTILEVAARAKS 420
 QY 570 FSRPPVTFKFMNDSTIY-----KQQLGV-----NNEILLTTTYLE----- 604
 DB 421 IRNKP-----QNNQLVFRKVLIKDLVDIERLKNLNAKKNKGVYLAESTYKELMDRVQ 475
 QY 605 -----TCQENTEY-----FQAKTD-----MYIYK-NYEHKTVPLSSITTLDTFIALNFTLL 651
 DB 476 NKDLLCOEQARKLEVLNKNVKSREQLQYVSKSQEHKKEVEALQQLVNSSTLESYKS 535
 QY 652 ENVDFFK---VTEL-----YTRDEKRLSNV-FDIETMFE--YNYVAQVSGVLRKDLLDLS 701
 DB 536 ENEKLKNELVLEIEKRYETNEAKITVATDLSQYRESKEYIASIYEKLDRTERNNKE 595
 QY 702 NRNOVFDAFGSLMDDDLGAQVGVVNAVSGVATL---FSSIVTGFINFIPKPFQGMGLMIIV 758
 DB 596 NENFNWLNKFNLLTMLRSFHGSPFTDNGYFTLLDNFNASMEELLNTHSN----- 645
 QY 759 VIGVLPALFYLTKKTKIYETAPIKMIYPEIDKLEREGKSEIAPISSEELERIVLAMHIH 818
 DB 646 -----QLLISMKTITEH-----FQSLDEALQSARSSCAVPNS--SLDLIVSEL--- 686
 QY 819 QONSHMETKTRKDPKDSILTRAQNMRLKRSYGS-NLKNAESVEML 862
 DB 687 -----KOSKSLDLALEHSLQDISSSQKUGNGISSELI 720

RESULT 39
 SYFB_MYCGE

FT	CONFLICT	81	94	NANNINPNPNKFP -> MLITITLIYHSCL (IN REF. 2).
FT	SEQUENCE	806 AA;	92132 MW;	D2F2BA6E9A064478 CRC64;
Qy	Query Match	3.0%;	Score 138.5;	DB 1; Length 806;
Qy	Best Local Similarity	19.6%;	Pred. No. 0.54;	
Qy	Matches 126;	Conservative	94;	Mismatches 201; Indels 221; Gaps 31;
Qy	114	SEGILLIYKQNIIPFMRVRKRYKRVVTTSTVYNGIYSDSIITNOHTTYKSIIEPWETEKMDT	173	
Qy	120	SEGILCSY-IDINPFS-----RQIIETEVA DAIIIDVSNDDHNDNQVLSFLSLD-DV	170	
Qy	174	IYOCFNSLRNLTGNNLLTYVDRODINWTFPLQVDPVGTDPDKRYGSGQPELYLBPFGWGS	233	
Qy	171	IF-----DVKTPTRNADLSLIFLAKELGVLLTKTFLKQ-----	205	
Qy	234	YRRRTTVNCELMDMFARSNPPDFEVTATGDTVEM-SPFWSGGEDHENKMKHEKPFVFS--	290	
Qy	206	--KSSVN-----HDFKFPFLNKKLKVANYPGGGLFQLQINQHSFPTVKGL	250	
Qy	291	VINN-YKVVDYQNRGTVPGLKTRIFLDREBYTILSWEKHLKMSYCPLTLWKAFYNGIQTE	349	
Qy	251	LINQMIKPVNY-----YVDK-----	265	
Qy	350	HSSGYHFVANDITASTTTSKEDMKENTTYHCLN-BEIKAEIEKKYAKVNSTHSHKYGDLK	408	
Qy	266	-----ANLVTV-FTAQ-----PIHCHDADRIVGNIELKQATHNETFVGLDDKQ	307	
Qy	409	YFKTDGGLYVWQPLIQNRLLDAKN-KLNNETYSRRSRQABESTTDPMMEMTGNAGGE	456	
Qy	308	YEIEPGDIV-----VCDEKGIALLVGIIGSKRTWVO-PTTNIFFEV-	349	
Qy	467	YSSENSITVAQVAVYDNIURINRIINTELDLSKAWCREQHRAALVWNELSKINPTSVMSMI	526	
Qy	350	--NCNSETIKQTA-----KRFLLNNF-----ASKPMVAKPISLLATDNCNLVYL	389	
Qy	527	YNRPVSAKRIGDIVSYSNCIIVD---QTSVSLHKSILRLLSASDEKCFSRPPTPKFMD	582	
Qy	390	QNSLLTTDNIIGKISHFSSSLKVFPEFKKUTVNPHKIRQLIG-----IEKKELTD	438	
Qy	583	STIVK--GQLG--VNNIEILLTTY---LETQCENTEYVFQAKTDMYIYKNYEHKTVPLS	635	
Qy	439	QTIKKSLSQLGFKVDNQLLKIPSYRODINTWQDISE-----EIVKLIDIN	483	
Qy	636	SITTLDTFIALNP-----TLLENVDF-KVIELYTRDEK--RLSNVFDIE	676	
Qy	484	KLKPIGITTSFNFSEKSSYFNTFNALTCLRKKLQTLGFHNIVITYQLTDQKSAKTNFLNLE	543	
Qy	677	TMFRFNYVYAAQVSGLRKDLDD---LSTN---RNQFVDAF	710	
Qy	544	NFTIKNPVSNHVMRVSILDSLLKVLKTNNNYKNELVNIF	585	
Qy	RESULT 40			
Qy	DYHC YEAST			
Qy	ID_DYHC YEAST	STANDARD;	PRT;	4092 AA.
Qy	AC P36022;			
Qy	DT 01-JUN-1994 (Rel. 29, Created)			
Qy	DT 01-JUN-1994 (Rel. 29, Last sequence update)			
Qy	DT 15-SEP-2003 (Rel. 42, Last annotation update)			
Qy	DE Dynein heavy chain, cytosolic (DYHC).			
Qy	GN DYN1 OR DHC1 OR YKR054C.			
Qy	OS Saccharomyces cerevisiae (Baker's yeast).			
Qy	OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;			
Qy	OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.			
Qy	NCBI_TaxID=4932;			
Qy	[1]			
Qy	RP SEQUENCE FROM N.A.			
Qy	RX MEDLINE=94068566; PubMed=8248224;			
Qy	RA Estel D., Urrestarazu L.A., Viissers S., Jauniaux J.-C.,			
Qy	RA van Vliet-Reedijk J.C., Planta R.J., Gibbons I.R.;			
Qy	RT "Cytoplasmic dynein is required for normal nuclear segregation in			
Qy	RT yeast.";			

```

RL Proc. Natl. Acad. Sci. U.S.A. 90:11172-11176(1993).
RN [2]
RP SEQUENCE FROM N.A.
RA Viessers S., Urtestarazu L.A., Jauniaux J.-C.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 1-3457 FROM N.A.
RX MEDLINE=94052110; PubMed=8234262;
RA Li Y.-Y., Yeh E.-Y., Hays T., Bloom K.S.;
RT "Disruption of mitotic spindle orientation in a yeast dynein mutant.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:10096-10100(1993).
CC -!- FUNCTION: Cytoplasmic dynein acts as a motor for the intracellular
CC retrograde motility of vesicles and organelles along microtubules.
CC Dynein has ATPase activity; the force-producing power stroke is
CC thought to occur on release of ADP. Required to maintain uniform
CC nuclear distribution in hyphae. May play an important role in the
CC proper orientation of the mitotic spindle into the budding
CC daughter cell yeast. Probably required for normal progression of
CC the cell cycle.
CC -!- SUBUNIT: Consists of at least two heavy chains and a number of
CC intermediate and light chains.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic. Probably binds indirectly to
CC the inner plasma membrane.
CC -!- DOMAIN: Dynein heavy chains probably consist of an N-terminal stem
CC (which binds cargo and interacts with other dynein components),
CC and the head or motor domain. The motor contains six tandemly-
CC linked AAA domains in the head, which form a ring. A stalk-like
CC structure (formed by two of the coiled coil domains) protrudes
CC between AAA 4 and AAA 5 and terminates in a microtubule-binding
CC site. A seventh domain may also contribute to this ring; it is not
CC clear whether the N-terminus or the C-terminus forms this extra
CC domain. There are four well-conserved and two non-conserved ATPase
CC sites, one per AAA domain. Probably only one of these (within AAA
CC 1) actually hydrolyzes ATP, the others may serve a regulatory
CC function.
CC -!- SIMILARITY: Belongs to the dynein heavy chain family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z21877; CAAT9923.1; -
CC DR EMBL; L15626; AAAL6055.1; -
CC DR EMBL; Z28279; CAAB2132.1; -
CC DR PIR; S38128; S38128.
CC DR SGD; S0001762; DYN1.
CC DR GO; GO:0005881; C:cytoplasmic microtubule; IDA.
CC DR GO; GO:0000132; P:mitotic spindle orientation; IGI.
CC DR InterPro; IPR003593; AAA_ATPase.
CC DR InterPro; IPR004273; Dynein_heavy.
CC DR Pfam; PF03028; Dynein_heavy; 1.
CC DR SMART; SM00382; AAA; 3.
CC DR Karyogamy; Motor protein; Dynein; Microtubules; ATP-binding; Repeat;
CC Coiled coil.
CC
CC FT DOMAIN 1 1757 STEM (BY SIMILARITY).
CC FT DOMAIN 1758 1979 AAA 1 (BY SIMILARITY).
CC FT DOMAIN 2036 2273 AAA 2 (BY SIMILARITY).
CC FT DOMAIN 2379 2628 AAA 3 (BY SIMILARITY).
CC FT DOMAIN 2722 2984 AAA 4 (BY SIMILARITY).
CC FT DOMAIN 2993 3300 STALK (BY SIMILARITY).
CC FT DOMAIN 3370 3599 AAA 5 (BY SIMILARITY).
CC FT DOMAIN 3760 3970 AAA 6 (BY SIMILARITY).
CC FT DOMAIN 154 175 COILED COIL (POTENTIAL).
CC FT DOMAIN 486 508 COILED COIL (POTENTIAL).
CC FT DOMAIN 542 566 COILED COIL (POTENTIAL).
CC FT DOMAIN 932 959 COILED COIL (POTENTIAL).
CC FT DOMAIN 1042 1063 COILED COIL (POTENTIAL).
CC FT DOMAIN 1681 1705 COILED COIL (POTENTIAL).
CC FT DOMAIN 2993 3092 COILED COIL (POTENTIAL).

```

```

FT DOMAIN 3242 3300 COILED COIL (POTENTIAL).
FT DOMAIN 3532 3608 COILED COIL (POTENTIAL).
FT NP_BIND 1796 1803 ATP (POTENTIAL).
FT NP_BIND 2074 2081 ATP (POTENTIAL).
FT NP_BIND 2418 2425 ATP (POTENTIAL).
FT NP_BIND 2760 2767 ATP (POTENTIAL).
FT CONFLICT 589 589 Y -> C (IN REF. 3).
FT CONFLICT 601 601 V -> A (IN REF. 3).
FT CONFLICT 1364 1364 E -> A (IN REF. 3).
FT CONFLICT 2118 2119 ML -> IV (IN REF. 1).
SQ SEQUENCE 4092 AA; 471337 MW; 3D9DF447E8E2D6BB CRC64;

Query Match
Best Local Similarity 3.0%; Score 134.5; DB 1; Length 4092;
Matches 119; Conservativity 103; Mismatches 234; Indels 141; Gaps 28;

QY 336 LTLKAFVNGIQTEHS---GSYHFVANDITASFTSKEDMKEFNTYHCLNEEIKAEIEK 392
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
441 LSTRKRFPFLAETIKSISPSYTHEEIQRLYHPF-----EQIHDSVNFRLKLEQAESEFSK 496
QY 393 KY-----AKVNSTHSGYGLKFKTDGGGLV-LV-WQPL-----IQNRLDAKKNLNNTYS 441
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
497 NMLDLERKLQNTLAFSFMDSHCHPTEKUSYLVKFKPLMELCPRIKVKVLENQOILLEI-- 554
QY 442 RRSRRQAEESTDPMEMTGAGGGEYSSENSITVAQVQAYD-NLRIRINNILEDLSKAW 500
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
555 KKDRIQLETGILELPKIL-----HVEALNIPISARISYFLNVQSIDNIV----- 601
QY 501 CREQHRAALV---WNELSKINPTSMYINRPVSAKRIGDIVSVNSCIVVDQTSVSLHK 557
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
602 --QYLEALFGSNWD-----TLEGR-----SISTSIQLRKETNPHD 636
QY 558 -SLRLLSASDEKCFSR---PPVTFKFMNDSTIYKGLGVNNEILLTITTYLTCQENTEY 613
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
637 VFLWLGNFPEKATANLLTTPILKLRNEDDY--ELKVNDFDAAAAYSE-LRSLTYMA 693
QY 614 FQ-----AKTDWYIKNVNHLKTVPLSSITLDTFI---ALNFTLLENVDFK--- 657
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
694 FQVPSHVIRIARTYMYLPR-----AINVELIQTFSSLSKLSYTFYNIPLKRNQV 746
QY 658 ---VIE--LYTRDEKRLSNVFDIETMPREYNYAQRVSGLRKOLLDLSTNRNQVDAFG 711
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
747 TWLLLLQILITPWSLEQESSEMSCSVHSLARLEKSIDGILSDYQILKNSEPOPAKEFS 806
QY 712 SI-----MDDLGAQGVTVNAVSGVATLFSIVTGFNFINKPFGGMLMIIVIGVLFA 765
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
807 GLKSGFDGTADDLHEVEEIIISNIOAFENLFTKGLTNVSDHI-STFNLLIISIIIEKVRN 865
QY 766 IYFLTTKTKIYETAPIKMIYPE-IDKLKEREGKSEIAP-----SEEELEIRVL 813
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
866 LK-----KMHFPKHVILKLSFNEGRITSSPSLAAMKRSLKLDIALLNKVVL 911
QY 814 AMHIHQNSHMETKTR-----KDPKDSILTBAQNMRLKRSYSLNKKNAESVEML 862
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
912 INFLHDPDHPPLSTLTFTNSLVILKDKDQNCIEQVQNLHCKINSY--VKEQKMEFL 966

RESULT 41
MLPI_YEAST STANDARD; PRT; 1875 AA.
AC Q02455;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin-like protein MLP1.
GN MLP1 OR YKR095W OR YKR415.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c;

```


233 SYR-----RRTTVNCELMDFARSPDPF---FVATGDTVMSP--WSGE--DDHEN 280
 548 SQKLSNNVENITLTSVEALGYSNKIYTFPLSLAEKWKGVQAGLPLNANVEVFTT 607
 281 KMHEKPF-----VSVINNY-----KVVDYQNRG-----TVP-L 308
 608 NIMKDTLDKISDVSVIPIGPAIINSGALRGNQAFATAGVAPLLSGFPFTIPAL 667
 309 GKTRIFLDREYTLSEKHLKNNSYC---PLTWKAPYNGIQTEHSGSYHFVANDITASF 365
 668 GVFTFYSSIOER-----EKIITKIENTCLEQRVKRWK-----DSYQMWVSNMLSRI 712
 366 TTSKEDMKEN-----TYYHCLN---BEIKAEIEKKYAKV-----NSTHS 402
 713 TT-----QFNHINYQMYDSLSVQADAKAKIDLEYKRYKSGSKENIKSOVENLKNSLDV 766
 403 KYGDL-----KYFKTDGGLVLMQPLIQNLQALDAKNNLNNETYSRRSRQAESTTDPME 457
 767 KISEAMNINKFIRECSVTY-----LFGNMLPKVIDELNK---FDLRTKTELINLID--- 815
 458 MTGNAGGEGYSSNITVAQVAYDNLRI-----NNILEDL----- 496
 816 -----SHNIIIVGEV-----DRLKAKVNESFENTMPPNIFSYTNNSLLKDIINEY 860
 497 -----SKANCREOHRALV-----WNELSKINPTSVMSIYNRPVSAKRIGD---VISVSN 544
 861 FNSINDSKILSLQKKNALVDTSGYNAEVRGVQNVQNTIYTNDFKLSSGDKIIVNLNN 920
 545 CI-----VVDQTSVSL-----HKSLRLLSASDEK-----CFSRPPVTFKFWNDST 584
 921 NILYSAIYENSVSFWIKISKDLTNSHNEYTIINSIQNSGKLCIRNGNIEWILQDVNR 980
 585 IYKQLGWNNEILLTYYLETQCNTEYYFOAKTD-----MYIKYVHEHKTVPLSIIT 639
 981 KYKSLIPDYSESHTGY-----TNKWFVFTTNNIMGYMKLYINGELKQSKIEDLDE 1034
 640 LDTFIALNFTLLENVDPKVIELYTRDEKLSNVFDIETMFEYN---YYAQRVSLRKO 695
 1035 VKLDKTVFGIDENIDENQMLWIRD-----FNIFSKELSNEDINIVYEGILRLNVKD 1087

RESULT 44
 DYHC_PARTE STANDARD; PRT; 4540 AA.
 AC Q27171;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Dynein heavy chain, cytosolic (DYHC).
 GN DYC-8.
 OS Paramesidium tetraurelia.
 OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Peniculida;
 OC Paramesidium.
 OC NCBI_TaxID=5888;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Stock 51;
 RX MEDLINE=96157890; PubMed=8589455;
 RA Kandel K.A., Forney J.D., Asai D.J.;
 RT "The dynein genes of Paramesidium tetraurelia: the structure and
 RT expression of the ciliary beta and cytoplasmic heavy chains.";
 RL Mol. Biol. Cell 6:1549-1562(1995).
 CC -!- FUNCTION: Cytoplasmic dynein acts as a motor for the intracellular
 CC retrograde motility of vesicles and organelles along microtubules.
 CC Dynein has ATPase activity; the force-producing power stroke is
 CC thought to occur on release of ADP.
 CC -!- SUBUNIT: Consists of at least two heavy chains and a number of
 CC intermediate and light chains.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- DOMAIN: Dynein heavy chains probably consist of an N-terminal stem
 CC (which binds cargo and interacts with other dynein components),
 CC and the head or motor domain. The motor contains six tandemly-

linked AAA domains in the head, which form a ring. A stalk-like
 structure (formed by two of the coiled coil domains) protrudes
 between AAA 4 and AAA 5 and terminates in a microtubule-binding
 site. A seventh domain may also contribute to this ring; it is not
 clear whether the N-terminus or the C-terminus forms this extra
 domain. There are four well-conserved and two non-conserved ATPase
 sites, one per AAA domain. Probably only one of these (within AAA
 3) actually hydrolyzes ATP, the others may serve a regulatory
 function.
 -!- SIMILARITY: Belongs to the dynein heavy chain family.
 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 EMBL; U20449; AAA75445.1; --
 PIR; T30838; T30838.
 InterPro; IPR003593; AAA ATPase.
 DR Pfam; PF03028; Dynein heavy; 1.
 DR SMART; SM00382; AAA; 3.
 KW Motor protein; Dynein; Microtubules; ATP-binding; Repeat; Coiled coil.
 FT DOMAIN 1 1796 STEM (BY SIMILARITY).
 FT DOMAIN 1797 2018 AAA 1 (BY SIMILARITY).
 FT DOMAIN 2091 2348 AAA 2 (BY SIMILARITY).
 FT DOMAIN 2091 2348 AAA 3 (BY SIMILARITY).
 FT DOMAIN 2796 3056 AAA 4 (BY SIMILARITY).
 FT DOMAIN 3076 3367 STALK (BY SIMILARITY).
 FT DOMAIN 3444 3673 AAA 5 (BY SIMILARITY).
 FT DOMAIN 3908 4123 AAA 6 (BY SIMILARITY).
 FT DOMAIN 440 482 COILED COIL (POTENTIAL).
 FT DOMAIN 698 722 COILED COIL (POTENTIAL).
 FT DOMAIN 794 827 COILED COIL (POTENTIAL).
 FT DOMAIN 975 995 COILED COIL (POTENTIAL).
 FT DOMAIN 1169 1251 COILED COIL (POTENTIAL).
 FT DOMAIN 1295 1311 COILED COIL (POTENTIAL).
 FT DOMAIN 3076 3182 COILED COIL (POTENTIAL).
 FT DOMAIN 3289 3367 COILED COIL (POTENTIAL).
 FT DOMAIN 3653 3688 COILED COIL (POTENTIAL).
 FT DOMAIN 3820 3851 COILED COIL (POTENTIAL).
 FT DOMAIN 4238 4259 COILED COIL (POTENTIAL).
 FT DOMAIN 4313 4342 COILED COIL (POTENTIAL).
 FT NP_BIND 1835 1842 ATP (POTENTIAL).
 FT NP_BIND 2129 2136 ATP (POTENTIAL).
 FT NP_BIND 2496 2503 ATP (POTENTIAL).
 FT NP_BIND 2834 2841 ATP (POTENTIAL).
 SQ SEQUENCE 4540 AA; 528626 MW; 0C6103148BFB95F9 CRC64;

Query Match 2.9%; Score 130; DB 1; Length 4540;
 Best Local Similarity 18.4%; Pred. No. 20;
 Matches 143; Conservative 130; Mismatches 290; Indels 216; Gaps 37;

QY 132 VRKYRKVTTSTVNGIYSDSITNQH-----TFYKSIPEWETEKMDTIYQCFNSLRLNTG 186
 Db 1021 IEKQQQLMDIKQGRATFDNFTTEHFGAIIIDYRMVQ-----VKINHKYDAWHKELLNH 1076
 QY 187 GNLLTVVDRDDINMTVFLQVDPGVDPVKRYGSOPELYLEPGWFGWGSYRRRTTVNCELMD 246
 Db 1077 GN-----KFGQLRVF-----NKNVTTEKELK 1100
 QY 247 MFARSNPPDPFFVTATGDTVE-----MSPFWSGEDDHENKMKHPFVSVINNYK 296
 Db 1101 I-----NFQDLTSDIIESITIIQEQDKKFPGWSADIE-----SFKNGOK 1139
 QY 297 VVD-----YQNRGTVPGLKTRIFLDREYTLSEKHLKNNSYCPLTWKAPYNGIQTEHSGS 353
 Db 1140 VLDRQRYQPGD-----WLSFEQVQWQWQ-FKQIRSKKLQSQSEMNNTQSIQQD 1190
 QY 354 YHVV---ANDITASFTTSKEDM-----KEFNNTTYHCLNEEIKAEIEKKYAKVNS----- 399

DR	PROSITE; PS00154; ATPASE_E1_E2; 1.	
KW	Hydrolase; Transmembrane; Phosphorylation; Magnesium; ATP-binding.	
FT	DOMAIN 1 35	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 36 58	POTENTIAL.
FT	DOMAIN 59 61	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 62 80	POTENTIAL.
FT	DOMAIN 81 407	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 408 427	POTENTIAL.
FT	DOMAIN 428 440	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 441 462	POTENTIAL.
FT	DOMAIN 463 1818	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 1819 1837	POTENTIAL.
FT	DOMAIN 1838 1845	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 1846 1863	POTENTIAL.
FT	DOMAIN 1864 1881	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 1882 1905	POTENTIAL.
FT	DOMAIN 1906 1928	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 1929 1952	POTENTIAL.
FT	DOMAIN 1953 1956	CYTOPLASMIC (POTENTIAL).
FT	MOD RES 496 496	PHOSPHORYLATION (BY SIMILARITY).
FT	METAL 1760 1760	MAGNESIUM (BY SIMILARITY).
FT	METAL 1764 1764	MAGNESIUM (BY SIMILARITY).
FT	DOMAIN 246 251	POLY-ASN.
FT	DOMAIN 252 256	POLY-LYS.
FT	DOMAIN 937 941	POLY-ASN.
FT	DOMAIN 1344 1347	POLY-LYS.
FT	DOMAIN 1363 1372	POLY-ASN.
FT	DOMAIN 1680 1684	POLY-ASN.
SQ	SEQUENCE 1956 AA; 230285 MW; AE708AAE99009335 CRC64;	
Query Match 2.8%; Score 129; DB 1; Length 1956;		
Best Local Similarity 18.0%; Pred. No. 7.1;		
Matches 134; Conservative 109; Mismatches 247; Indels 254; Gaps 34;		

[illegible]

Qy 504 OHRAALVWNELSKINPTSVMSMIYNRPVSAKRIGDVISVSNCIWVDTSVSLHKSLELLS 563
Db 1442 -----YIYEKKYLOHCLLKHNDNYKKVLPRIKD-INYSYQMESIKTRNFH-----S 1489
Qy 564 ASDEKCFSPRPVTFKFW-----NDSTIYKGOLGVNNEILLTTTYLETQCENTEY--- 612
Db 1490 LSEOFAPSNLILSFYIIKNDNDNNVYKNQYIYNKYIYNKNSICNKNVI--CNKNYIYNKN 1547
Qy 613 YFOAKTDMYIYKN-YEHLKTVPLS 635
Db 1548 NIYNKNYIYNKNILTHAKSVLLS 1571

Search completed: October 28, 2003, 15:31:34
Job time : 38 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 15:27:11 ; Search time 118 Seconds
(without alignments)
1891.657 Million cell updates/sec

Title: US-10-055-364-24

Perfect score: 4547

Sequence: 1 MAGSLKRGSLALWLYQV.....KRGYSNLKNAESVEMLNTL 865

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23.*

1: sp_archaea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mhc.*

8: sp_organelle.*

9: sp_phage.*

10: sp_plant.*

11: sp_rodent.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

15: sp_rvirus.*

16: sp_bacteriap.*

17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	4547	100.0	865 12 Q913V4	Q913V4 porcine gam
2	4547	100.0	876 12 Q8JYD3	Q8JYD3 porcine lym
3	4271	93.9	876 12 Q8B3X4	Q8B3X4 porcine lym
4	4265	93.8	876 12 Q8B423	Q8B423 porcine lym
5	3529	77.6	876 12 Q8AYX3	Q8AYX3 porcine lym
6	2513	55.3	863 12 Q91A92	Q91A92 ovine herpe
7	2504	55.1	863 12 Q91A90	Q91A90 ovine herpe
8	2493.5	54.8	863 12 Q91A89	Q91A89 ovine herpe
9	2491	54.8	863 12 Q91A91	Q91A91 ovine herpe
10	2489.5	54.8	854 12 Q36362	Q36362 adcelaphine
11	1962.5	43.2	815 12 Q8JUX2	Q8JUX2 badger herp
12	1953.5	43.0	874 12 Q86613	Q86613 equine herp
13	1937	42.6	845 12 P88906	P88906 kaposi's sa
14	1927.5	42.4	829 12 Q9WR79	Q9WR79 macaca mula
15	1927.5	42.4	829 12 Q39233	Q39233 thesus monk
16	1925.5	42.3	829 12 Q9J2M4	Q9J2M4 macaca mula

ALIGNMENTS

RESULT 1

```

Q913V4
ID Q913V4 PRELIMINARY; PRT; 865 AA.
AC Q913V4;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE Envelope glycoprotein B.
GN GPB.
OS Porcine gamma-herpesvirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae.
OX NCBI_TaxID=169968;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21124355;
RA Huang C., Fuchimoto Y., Gleit Z., Ericsson T., Greissner A.,
RA Schier-Dolberg R., Melendy E., Kitamura H., Fishman J., Ferry J.,
RA Harris N., Patience C., Sachs D.;
RT "Posttransplantation lymphoproliferative disease in miniature swine
RT after allogeneic hematopoietic cell transplantation: similarity to
RT human PTLD and association with a porcine gammaherpesvirus.";
RL Blood 97:1467-1473(2001).
DR EMBL; AY044427; AAK95853.1; -.
DR InterPro; IPR00234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
SQ SEQUENCE 865 AA; 99291 MW; E3F6CEFI1B3EB326E CRC64;

```

Query Match 100.0%; Score 4547; DB 12; Length 865;
Best Local Similarity 100.0%; Pred. NO. 1.1e-277;
Matches 865; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAGSLKRGSLALWLYQVALYSLSTAETGVTSPPNTATWSTESPLTGYTHDSHGE 60

Db 1 MAGSLKRGSLALWLYQVALYSLSTAETGVTSPPNTATWSTESPLTGYTHDSHGE 60

Qy 61 RGNENRDSSEONKNVIGSPSTFPYRVCSAGVGDFRFTDHDVCPDASDMVHSEGI 120

Q9YTD5 ateline her
Q72004 equine herp
Q993K7 callitrichi
Q8UZD5 cercopithi
Q98336 murid herpe
Q65561 bovine herp
Q913V5 baboon gamt
Q99BV2 baboon cyto
Q96876 human cytom
Q89172 human cytom
Q85432 cercopithe
Q99169 human cytom
Q69170 human cytom
Q69173 human cytom
Q992K6 ovine herpe
Q89163 human cytom
Q992K7 bovine lym
Q39989 human herpe
Q8V615 porcine cyt
Q96910 human herpe
Q56280 human herpe
Q98VQ0 caprine her
Q8V429 porcine cyt
Q8V429 asian eleph
Q8V430 porcine cyt
Q8V614 porcine cyt
Q8V613 porcine cyt
Q8V613 human herpe
Q8Q833 chimpanzee

17 1925 42.3 807 12 Q9YTD5
18 1898.5 41.8 870 12 Q72004
19 1841 40.5 870 12 Q993K7
20 1812 39.9 864 12 Q8UZD5
21 1796 39.5 849 12 Q98336
22 1787 39.3 874 12 Q65561
23 1656 36.4 752 12 Q913V5
24 1159.5 25.5 877 12 Q99BV2
25 1131 24.9 907 12 Q96876
26 1129 24.8 907 12 Q89172
27 1126 24.8 854 12 Q85432
28 1123.5 24.7 906 12 Q99169
29 1121.5 24.7 906 12 Q69170
30 1119 24.6 905 12 Q69173
31 1117 24.6 352 12 Q992K6
32 1110 24.4 905 12 Q89163
33 1097 24.1 340 12 Q992K7
34 1090.5 24.0 822 12 Q39989
35 1087.5 23.9 859 12 Q8V615
36 1086.5 23.9 822 12 Q96910
37 1086.5 23.9 822 12 Q56280
38 1084.5 23.9 349 12 Q98VQ0
39 1080.5 23.8 851 12 Q8V429
40 1071 23.6 845 12 Q9DKT9
41 1063 23.4 860 12 Q8V430
42 1061 23.3 859 12 Q8V614
43 1059 23.3 860 12 Q8V613
44 1052.5 23.1 830 12 Q9W125
45 1047.5 23.0 921 12 Q8Q833

```

Db 61  RGNENRDESEONKNIYGSPTFPYRVCSASGVGVDFRQTDHVCPCDASDMVHSEGIILLI 120
Qy 121  YKQNIIPFMRVRKRVKVVTTSTVYNGIYSDSITNQHTFYKSIETPWTETKMDTIYQCFNS 180
Db 121  YKQNIIPFMRVRKRVKVVTTSTVYNGIYSDSITNQHTFYKSIETPWTETKMDTIYQCFNS 180
Qy 181  LRLNTGNNLTIVYDRIINNTVFLQPDVGVTPDKRYGSOPELYLEPGWFGSVYRRRTTV 240
Db 181  LRLNTGNNLTIVYDRIINNTVFLQPDVGVTPDKRYGSOPELYLEPGWFGSVYRRRTTV 240
Qy 241  NCELMDMFARSNPPDFVTATGDTVEMSPFWSGEDDHENKMKHEKPFVSVINNYKVVDY 300
Db 241  NCELMDMFARSNPPDFVTATGDTVEMSPFWSGEDDHENKMKHEKPFVSVINNYKVVDY 300
Qy 301  QNRGTVPGLKTRIFLDREBYTLSWEKHLKMSYCPULTLWKAIFYNGIQTSHSGSYHFVAND 360
Db 301  QNRGTVPGLKTRIFLDREBYTLSWEKHLKMSYCPULTLWKAIFYNGIQTSHSGSYHFVAND 360
Qy 361  ITASFTTSKEDMKEFNTTYHCLNEEIKAEIEKKYAKVNSTHSGYKGDLYKFKTDGGLYLW 420
Db 361  ITASFTTSKEDMKEFNTTYHCLNEEIKAEIEKKYAKVNSTHSGYKGDLYKFKTDGGLYLW 420
Qy 421  QPLIQNRLLDKAKNNETYSRRRQAESTTDPMMEMTNGAGGEYSSENSITVAQVOY 480
Db 421  QPLIQNRLLDKAKNNETYSRRRQAESTTDPMMEMTNGAGGEYSSENSITVAQVOY 480
Qy 481  AYDNLRIIRINNILEDLSKAWCREQRAALVWNLKSKINPTSVMSMIYNRPVSAKRIGDVI 540
Db 481  AYDNLRIIRINNILEDLSKAWCREQRAALVWNLKSKINPTSVMSMIYNRPVSAKRIGDVI 540
Qy 541  SVSNCIYVDQTSVSLHSLRLLSASDEKCFSRPPVTFKFMNDSTIYKQGLGVNNEILLTT 600
Db 541  SVSNCIYVDQTSVSLHSLRLLSASDEKCFSRPPVTFKFMNDSTIYKQGLGVNNEILLTT 600
Qy 601  TYLETQCENTEYFFQAQTDWYIYKNYEHLKTVPLSSITTLDTFIALNFTLLENVDFKVE 660
Db 601  TYLETQCENTEYFFQAQTDWYIYKNYEHLKTVPLSSITTLDTFIALNFTLLENVDFKVE 660
Qy 661  LYTRDEKRLSNVDFIETMFREYNYAQRVSGRLKDLDDLSTNRNQFVDAFGSLMDDLGA 720
Db 661  LYTRDEKRLSNVDFIETMFREYNYAQRVSGRLKDLDDLSTNRNQFVDAFGSLMDDLGA 720
Qy 721  GOTVNAVSGVATLFFSSIVTGFINFKNPFGGMLIIVLGVLPFAIYFLTKTKIYETAP 780
Db 721  GOTVNAVSGVATLFFSSIVTGFINFKNPFGGMLIIVLGVLPFAIYFLTKTKIYETAP 780
Qy 781  IKMIYPIDKLKREGKSEIAPISEEELERIVLAMHIHQONSHMETKTRDKPXSILTRA 840
Db 781  IKMIYPIDKLKREGKSEIAPISEEELERIVLAMHIHQONSHMETKTRDKPXSILTRA 840
Qy 841  QNMLRKRSYGNLKNABSEVEMNLTL 865
Db 841  QNMLRKRSYGNLKNABSEVEMNLTL 865

```

RESULT 2

```

Q8UYD3 ID Q8UYD3 PRELIMINARY; PRT; 876 AA.
AC Q8UYD3;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Glycoprotein B.
OS Porcine lymphotropic herpesvirus 1.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae.
OX NCBI_TaxID=91740;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=sample #56;
RX MEDLINE=99226949; PubMed=10211967;
RA Ehlers B., Ulrich S., Goltz M.;
RT "Detection of two novel porcine herpesviruses with high similarity to

```

```

RT gammaherpesviruses.";
RL J. Gen. Virol. 80:971-978(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=sample #56;
RX MEDLINE=20036635; PubMed=10567652;
RA Ulrich S., Goltz M., Ehlers B.;
RT "Characterization of the DNA polymerase loci of the novel porcine
RT lymphotropic herpesviruses 1 and 2 in domestic and feral pigs.";
RL J. Gen. Virol. 80:3199-3205(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=sample #56;
RX MEDLINE=22008485; PubMed=12009880;
RA Goltz M., Ericsson T., Patience C., Huang C.A., Noack S., Sachs D.H.,
RA Ehlers B.;
RT "Sequence analysis of the genome of porcine lymphotropic herpesvirus 1
RT and gene expression during post-transplant lymphoproliferative disease
RT of pigs.";
RL Virology 294:393-393(2002).
DR EMBL; AF478169; AAM22114.1; -.
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
SQ SEQUENCE 876 AA; 100705 MW; 2D1D4DB61F5C0D24 CRC64;

```

```

Query Match 100.0%; Score 4547; DB 12; Length 876;
Best Local Similarity 100.0%; Pred. No. 1.1e-277;
Matches 865; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1  MAGSLKRGSLALWLYQVALYSIAETGVTSPENTATWSTESPLTGHYTHDSHGE 60
Db 12  MAGSLKRGSLALWLYQVALYSIAETGVTSPENTATWSTESPLTGHYTHDSHGE 71
Qy 61  RGNENRDESEONKNIYGSPTFPYRVCSASGVGVDFRQTDHVCPCDASDMVHSEGIILLI 120
Db 72  RGNENRDESEONKNIYGSPTFPYRVCSASGVGVDFRQTDHVCPCDASDMVHSEGIILLI 131
Qy 121  YKQNIIPFMRVRKRVKVVTTSTVYNGIYSDSITNQHTFYKSIETPWTETKMDTIYQCFNS 180
Db 132  YKQNIIPFMRVRKRVKVVTTSTVYNGIYSDSITNQHTFYKSIETPWTETKMDTIYQCFNS 191
Qy 181  LRLNTGNNLTIVYDRIINNTVFLQPDVGVTPDKRYGSOPELYLEPGWFGSVYRRRTTV 240
Db 192  LRLNTGNNLTIVYDRIINNTVFLQPDVGVTPDKRYGSOPELYLEPGWFGSVYRRRTTV 251
Qy 241  NCELMDMFARSNPPDFVTATGDTVEMSPFWSGEDDHENKMKHEKPFVSVINNYKVVDY 300
Db 252  NCELMDMFARSNPPDFVTATGDTVEMSPFWSGEDDHENKMKHEKPFVSVINNYKVVDY 311
Qy 301  QNRGTVPGLKTRIFLDREBYTLSWEKHLKMSYCPULTLWKAIFYNGIQTSHSGSYHFVAND 360
Db 312  QNRGTVPGLKTRIFLDREBYTLSWEKHLKMSYCPULTLWKAIFYNGIQTSHSGSYHFVAND 371
Qy 361  ITASFTTSKEDMKEFNTTYHCLNEEIKAEIEKKYAKVNSTHSGYKGDLYKFKTDGGLYLW 420
Db 372  ITASFTTSKEDMKEFNTTYHCLNEEIKAEIEKKYAKVNSTHSGYKGDLYKFKTDGGLYLW 431
Qy 421  QPLIQNRLLDKAKNNETYSRRRQAESTTDPMMEMTNGAGGEYSSENSITVAQVOY 480
Db 432  QPLIQNRLLDKAKNNETYSRRRQAESTTDPMMEMTNGAGGEYSSENSITVAQVOY 491
Qy 481  AYDNLRIIRINNILEDLSKAWCREQRAALVWNLKSKINPTSVMSMIYNRPVSAKRIGDVI 540
Db 492  AYDNLRIIRINNILEDLSKAWCREQRAALVWNLKSKINPTSVMSMIYNRPVSAKRIGDVI 551
Qy 541  SVSNCIYVDQTSVSLHSLRLLSASDEKCFSRPPVTFKFMNDSTIYKQGLGVNNEILLTT 600
Db 552  SVSNCIYVDQTSVSLHSLRLLSASDEKCFSRPPVTFKFMNDSTIYKQGLGVNNEILLTT 611
Qy 601  TYLETQCENTEYFFQAQTDWYIYKNYEHLKTVPLSSITTLDTFIALNFTLLENVDFKVE 660
Db 612  TYLETQCENTEYFFQAQTDWYIYKNYEHLKTVPLSSITTLDTFIALNFTLLENVDFKVE 671

```

```
QY 661 LYTRDEKRLSNVFDIETMFREYNYAQRVSGLRKOLLDLSTNRNPFVDAFGSLMDDLGA 720
Db 672 LYTRDEKRLSNVFDIETMFREYNYAQRVSGLRKOLLDLSTNRNPFVDAFGSLMDDLGA 731
QY 721 GQTVNAVSGVATLSSIVTGFINFKNPFQGMMLIIIVIGVLFYFELTKTKIYETAP 780
Db 732 GQTVNAVSGVATLSSIVTGFINFKNPFQGMMLIIIVIGVLFYFELTKTKIYETAP 791
QY 781 IKMIYPEIDKLKREGKSEIAPISSEELERIVLAMHIHQONSHMETKTRKPKDSILTRA 840
Db 792 IKMIYPEIDKLKREGKSEIAPISSEELERIVLAMHIHQONSHMETKTRKPKDSILTRA 851
QY 841 QNMLRKSGSYNLKNAESVEMLNTL 865
Db 852 QNMLRKSGSYNLKNAESVEMLNTL 876

RESULT 3
Q8B3X4
ID Q8B3X4 PRELIMINARY; PRT; 876 AA.
AC Q8B3X4;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Glycoprotein B.
OS Porcine lymphotropic herpesvirus 2.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae.
OX NCBI_TaxID=91741;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=568;
RA Chmielewicz B., Goltz M., Franz T., Bauer C., Brema S., Ellerbrok H.,
RA Beckmann S., Rziha H.-J., Lahrmann K.-H., Romero C., Ehlers B.;
RA "A novel porcine gammaherpesvirus.";
RL Virology 0:0-0(2003).
DR EMBL; AY170317; AA012353.1; -.
SQ SEQUENCE 876 AA; 100741 MW; 4CB870F431BCB0DF CRC64;

Query Match 93.9%; Score 4271; DB 12; Length 876;
Best Local Similarity 93.4%; Pred. No. 2.5e-260;
Matches 808; Conservative 28; Mismatches 29; Indels 0; Gaps 0;

QY 1 MAGSLKRGSLALWLYQVALYSLSIATGTGVTSPNTATWSTESPLTGHYTHDSSHG 60
Db 12 MAGSLKRLRALWCQFQVALYSLKAEKATSANTQAWSTEALPIGQYGYDSSHG 71

QY 61 RGNENRDSSEQNKNIYGSPTFPYRVCSAGVGDFRQTDHVCDDASDMVHSEGI 120
Db 72 RATSSENDEEHNKNIYGSPTFPYRVCSAGVGDFRQTDHVCDDASDMVHSEGI 131

QY 121 YKONIIPFMRVRKRVKVTSTVNGIYSDSIHQHTFYKSIPEWTEKMDTIYQC 180
Db 132 YKONIIPFMRVRKRVKVTSTVNGIYSDSIHQHTFYKSIPEWTEKMDTIYQC 191

QY 181 LRLNTGGNLLTYVDRDDINMTVFLQPDVGVTPDKRYGSOPELYLEPGWFGSYRRRT 240
Db 192 LRLNTGGNLLTYVDRDDINMTVFLQPDVGVTPDKRYGSOPELYLEPGWFGSYRRRT 251

QY 241 NCELMDFARNSPPDFVTATGTVEMSPFWSGEDDHENKHKPFVSPVNNYKVVDY 300
Db 252 NCELMDFARNSPPDFVTATGTVEMSPFWSGEDDHENKHKPFVSPVNNYKVVDY 311

QY 301 QNRGTVPLGKTRIFLDREYTLWSWKHLKNWSYCPPLTWKAFYNGIQTEHSGSYH 360
Db 312 QNRGTVPLGKTRIFLDREYTLWSWKHLKNWSYCPPLTWKAFYNGIQTEHSGSYH 371

QY 361 ITASFTTSKEDMKFNTYHCLNEEIKAEIEKKYAKVNSTHYSKYGDLYFKTDGGL 420
Db 372 ITASFTTSKEDMKFNTYHCLNEEIKAEIEKKYAKVNSTHYSKYGDLYFKTDGGL 431

QY 421 OPLIQNRLLDKAKNLNNETYSRRSRQAESTTDPMMETGNGAGGEYSSENSITVAQ 480
```

```
Db 432 QPLIQNRLLDKAKNLNNETYSRRSRQVDSITGPMMEATGNGAGGEYSSENSITVAQ 491
QY 481 AYDNLRTIRNNILBEDLSKAWCREQHRALVWNELSKINPTSVMSMIYNRPVSAKRIG 540
Db 492 AYDNLRTIRNNILBEDLSKAWCREQHRALVWNELSKINPTSVMSMIYNRPVSAKRIG 551
QY 541 SVSNCIVVDQTSVSLHKSRLLSASDEKCFRPPVTFKFMNDSTIYKQGLGVNNEILL 600
Db 552 SVSNCIVVDQTSVSLHKSRLLSASDEKCFRPPVTFKFMNDSTIYKQGLGVNNEILL 611
QY 601 TYLETCQENTEYFYQAOKTDMYIYKNYEHKLTVPPLSSITTLDTFIALNFTLENVDF 660
Db 612 TYLETCQENTEYFYQAOKTDMYIYKNYEHKLTVPPLSSITTLDTFIALNFTLENVDF 671
QY 661 LYTRDEKRLSNVFDIETMFREYNYAQRVSGLRKOLLDLSTNRNPFVDAFGSLMDDLGA 720
Db 672 LYTRDEKRLSNVFDIETMFREYNYAQRVSGLRKOLLDLSTNRNPFVDAFGSLMDDLGA 731
QY 721 GQTVNAVSGVATLSSIVTGFINFKNPFQGMMLIIIVIGVLFYFELTKTKIYETAP 780
Db 732 GQTVNAVSGVATLSSIVTGFINFKNPFQGMMLIIIVIGVLFYFELTKTKIYETAP 791
QY 781 IKMIYPEIDKLKREGKSEIAPISSEELERIVLAMHIHQONSHMETKTRKPKDSILTRA 840
Db 792 IKMIYPEIDKLKREGKSEIAPISSEELERIVLAMHIHQONSHMETKTRKPKDSILTRA 851
QY 841 QNMLRKSGSYNLKNAESVEMLNTL 865
Db 852 QNMLRKSGSYNLKNAESVEMLNTL 876

RESULT 4
Q8B423
ID Q8B423 PRELIMINARY; PRT; 876 AA.
AC Q8B423;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Glycoprotein B.
OS Porcine lymphotropic herpesvirus 2.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae.
OX NCBI_TaxID=91741;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=546;
RA Chmielewicz B., Goltz M., Franz T., Bauer C., Brema S., Ellerbrok H.,
RA Beckmann S., Rziha H.-J., Lahrmann K.-H., Romero C., Ehlers B.;
RA "A novel porcine gammaherpesvirus.";
RL Virology 0:0-0(2003).
DR EMBL; AY170314; AA012281.1; -.
SQ SEQUENCE 876 AA; 100771 MW; 3B238639221474C7 CRC64;

Query Match 93.8%; Score 4265; DB 12; Length 876;
Best Local Similarity 93.3%; Pred. No. 6e-260;
Matches 807; Conservative 28; Mismatches 30; Indels 0; Gaps 0;

QY 1 MAGSLKRGSLALWLYQVALYSLSIATGTGVTSPNTATWSTESPLTGHYTHDSSHG 60
Db 12 MAGSLKRLRALWCQFQVALYSLKAEKATSANTQAWSTEALPIGQYGYDSSHG 71

QY 61 RGNENRDSSEQNKNIYGSPTFPYRVCSAGVGDFRQTDHVCDDASDMVHSEGI 120
Db 72 RATSSENDEEHNKNIYGSPTFPYRVCSAGVGDFRQTDHVCDDASDMVHSEGI 131

QY 121 YKONIIPFMRVRKRVKVTSTVNGIYSDSIHQHTFYKSIPEWTEKMDTIYQC 180
Db 132 YKONIIPFMRVRKRVKVTSTVNGIYSDSIHQHTFYKSIPEWTEKMDTIYQC 191

QY 181 LRLNTGGNLLTYVDRDDINMTVFLQPDVGVTPDKRYGSOPELYLEPGWFGSYRRRT 240
Db 192 LRLNTGGNLLTYVDRDDINMTVFLQPDVGVTPDKRYGSOPELYLEPGWFGSYRRRT 251
```



```

QY 555 LKSLRLLSASD-EKCFSPRPVTFKFMNDSTIYKGLQGVNNEILLTTTTLTETCOENTYY 613
D 545 LHRNRPGRDHAHECVSRPPTVKFINDSHLYKGLQGVNNEILLTTTTLTALVCHENTHY 604
QY 614 FOAKTDMYIYKYNVYKHTVPLUSSITTLDTFIALNFTLLENVDFKVIELYTRDEKRLSNVF 673
D 605 FQGGNNYFYKYNVYKHTVPIVSAVATLDTFVLNLTIVENIDFQVIELYSREKRMSTVF 664
QY 674 DIETMREYNYAQRVSLGRKDLDDLSTNRNOFVDAFGLMDDLGAQGVTVNAVSVAT 733
D 665 DIETMREYNYTORVGLRDLSDIATNRNQFVDAFGTMDLGGVGVKTVNAVSSLAT 724
QY 734 LFSSIVTGFINFKNPGFGLMIMIVGVLFAIYFLTKTKIYETAPIKMIYPEIDKLE 793
D 725 LFSSIVTGFINFKNPGFGLMIFGLLAUVIALLRRRAASFAANVQMIYDIOQITK 784
QY 794 REGKEIAPISBEELERIVLAMIHQONSHMETKTRKDP-----KDSILTRAQNMRLK 846
D 785 QROEMNVEPIKSHELDRIMLAMHDYHQTQDKPDEKEGPESGGSANKANLNKAKNVLRR 844
QY 847 RSGVSNLKNNAESVE 860
D 845 RAGYQPLKRSDDSTE 858

RESULT 8
Q91A89.
ID Q91A89 PRELIMINARY; PRT; 863 AA.
AC Q91A89;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Glycoprotein B.
GN GB.
OS Ovine herpesvirus 2.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae.
OX NCBI_TaxID=10398;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21488530; PubMed=11602790;
RA Dunowska M., Letchworth G.J., Collins J.K., DeMartini J.C.;
RT "Ovine herpesvirus-2 glycoprotein B sequences from tissues of ruminant
RT malignant catarrhal fever cases and healthy sheep are highly
RT conserved.";
RL J. Gen. Virol. 82:2785-2790(2001).
DR EMBL; AF385442; AAL24472.1; -.
DR InterPro; IPR000234; Glycoprot B.
DR Pfam; PF00606; Glycoprotein B_1.
DR ProDom; PD000693; Glycoprot_B_1.
SQ SEQUENCE 863 AA; 97534 MW; B2F63AB88876A777 CRC64;

Query Match 54.8%; Score 2493.5; DB 12; Length 863;
Best Local Similarity 55.4%; Pred. No. 1.8e-148;
Matches 475; Conservative 149; Mismatches 213; Indels 21; Gaps 5;

QY 17 LYQVALYSLSTAETAGVTSPTATWSTESPLTGHYGHDSHGEGNENRDSSEQKNVI 76
D 8 LSPAAITLLCLAAACLGQTPTPTPTEDDVIYVGNVSPDVLKQOLGSLNEEGSINRGP 67
QY 77 YGSPSTFPYVCSAGVGDFRFTDHCVPDASDMVHSEGILLITYKONIIPFMFRVKYR 136
D 68 M-DPSAFPYVCSASNIGDIFRQTHSCNTKDKENEGILLIFKENIYVYFVKYR 126
QY 137 KVVTSTVYNGISDSTNQHFTFKYSIEPWETKMDTIYQCFNSRLNTCGNLLTYVD RD 196
D 127 KIVTSTVYNGISDAITNQHFSKSPHYEARMDTIYQCVNSLSLTGCGNLLAYTDND 186
QY 197 DINMTVFLQVGVTPDVVKYSGOPELYLEPGHFWGSYRRRTTNCELMDMFARPPPD 256
D 187 GYNLTVDLQPMGLNSVRRNSVQPEIHAEPGRLLGGYRRRTTNCVETDTEARSVPFPR 246

```

```

QY 257 FFVTATGDTIVEMSPFWSGEDDHENKHEKPFWFVSVINNVKVDYQNRGTVPGLKTRIFLD 316
D 247 YFVTNVGDITIEKSPFWSGGANETEPNKRRTVSVLRDYLTVDYKGRSPQPHTRIFID 306
QY 317 REBYTILSWKHLKMSYCPILTLWKAFYNGIQTEHSGSYHFVANDITASFTTSKEDMKBFN 376
D 307 KEDYTLSSWAQQLKNISYCRWAHWSFHNAIKTEHENSYHFVANDITASFTTPTEAQDVT 366
QY 377 TTVHCLNEEBIKABIEKKYAKVNSTHSGYGLDKYFKTDGGLYLVMOPLIQNRLLDANKL- 435
D 367 KTHCLNSLIESBMTSRLEKVGTHVNGSRQYLLTNGGLLVQPLVQOQKLLNAQDLLE 426
QY 436 -----NNETYSRRSRQAESTTDPMMEMTCNGAGGYSSENSITVAQOVAYDNLIRIN 490
D 427 AVASKENVTESARSRRQRAVSSILID-----DDVYTAESALLTQIQFAYDMLRSQIN 480
QY 491 NILEDISKAWCREQHRALVWNELSKINPTSVMSIMYNRPVSAKRIGDVISVSNICIVDQ 550
D 481 TVLEELSRACRQHRASLWMNELSKINPTSVMSIYGRPVSAKRIGDVISVSHCVVDQ 540
QY 551 TSVSLHKSRLLSASD-EKCFSPRPVTFKFMNDSTIYKGLQGVNNEILLTTTTLTETCOEN 609
D 541 QSVSLHRSRMKVPGRDHAHECVSRPPTVKFINDSHLYKGLQGVNNEILLTTTTLTALVCHEN 600
QY 610 TEYFQAKTDMYIYKYNVYKHTVPLSSITTLDTFFIALNFTLLENVDFKVIELYTRDEKRL 669
D 601 TEHYFOGNNYFYKYNVYKHTVPIVSAVATLDTFVLNLTIVENIDFQVIELYSREKRM 660
QY 670 SNVFDIETMREYNYAQRVSLGRKDLDDLSTNRNOFVDAFGLMDDLGAQGVTVNAVS 729
D 661 STVFDIETMREYNYTORVGLRDLSDIATNRNQFVDAFGTMDLGGVGVKTVNAVS 720
QY 730 GYATLFSSIVTGFINFKNPGFGLMIMIVGVLFAIYFLTKTKIYETAPIKMIYPEID 789
D 721 SLATLFSSIVTGFINFKNPGFGLIFGLAAVIAVILLRRRAASFAANVQMIYDPTIQ 780
QY 790 KLKEREKGEIAPISBEELERIVLAMIHQONSHMETKTRKDP-----KDSILTRAQN 842
D 781 QITKQEQMNVEPIKSHELDRIMLAMHDYHQTQDKPDEKEGPESGGSANKANLNKAKN 840
QY 843 MLRKSGVSNLKNNAESVE 860
D 841 VLRRRAGYQPLKRSDDSTE 858

RESULT 9
Q91A91
ID Q91A91 PRELIMINARY; PRT; 863 AA.
AC Q91A91;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Glycoprotein B.
GN GB.
OS Ovine herpesvirus 2.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae.
OX NCBI_TaxID=10398;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21488530; PubMed=11602790;
RA Dunowska M., Letchworth G.J., Collins J.K., DeMartini J.C.;
RT "Ovine herpesvirus-2 glycoprotein B sequences from tissues of ruminant
RT malignant catarrhal fever cases and healthy sheep are highly
RT conserved.";
RL J. Gen. Virol. 82:2785-2790(2001).
DR EMBL; AF385440; AAL24470.1; -.
DR InterPro; IPR000234; Glycoprot B.
DR Pfam; PF00606; Glycoprotein B_1.
DR ProDom; PD000693; Glycoprot_B_1.
SQ SEQUENCE 863 AA; 97512 MW; 9464D7BD6DB863C1 CRC64;

Query Match 54.8%; Score 2491; DB 12; Length 863;

```

Best Local Similarity 55.6%; Pred. No. 2.5e-148;		Matches 473; Conservative 145; Mismatches 199; Indels 34; Gaps 7;	
QY	24	SLSIAGTGTSPNTATWTEGSLTGHYCTHDSHGERGNNENRDSEONKNIYGSPTF	83
DB	28	TLPTTDDVIVGNTV-----SP-----DVLKQOLGSGUNESEGSINRGM-DPSAF	73
QY	84	PYRVCSAGVGDFRQTDHVCPCDASDMVHSEGLIYKQNTLIPFMFRVKYRKVVTTST	143
DB	74	PRVCSASNIGDIFRQTHSCPTNKDKEHNEGILLIPKENIVPVYFKYRKIVTTST	133
QY	144	VYNGIYSDITNOHTFYKISBPWETEKMDTIYQCFNSRLNTGCGNLLTYVDRDDINMTVF	203
DB	134	VYNGIYSDAITNOHESFSPHYEARMDTIYQCFNSLSLTVGGNLLAYTNDGYNLATVD	193
QY	204	LQPVGVTPDVKRYGQPELYLEPGWFGSYRRRTTVNCELMDMFARSPDPDFVTTATG	263
DB	194	LQPMGLNSVRVYNSQPEIHAEPOWLLGGYRRRTTVNCEVDTTARSVPPPIRYVTWG	253
QY	264	DTVEMSPFWSGDDHDKHKKHKEPWFVSVINNVKVDYQNRGTVPGLKTRIFLDRBEYTL	323
DB	254	DTIEMSPFWSGCGANETEPNKEPRRTVSVLRYDTLVYDKDGRSPQPHTRIFIDKEDYTL	313
QY	324	WEKHLKMSYCYPLTWKAPYNGIQTEHSGSYHFVANDITASPTTSKEDMKENFTTYHCLN	383
DB	314	MAQOLKNISYCRWAHWSFHNAIKTEHNSYHFVANDITASPTTNTAQAQVTKTHTCLN	373
QY	384	EIKAEIEKKYAKVNSTHSGYGLDYKFTDGGYLVWQPLQNLNLDLDAKNKL-----NN	437
DB	374	SLIESEMTSRLEKVGTHVNGRSQYLTNGGLLVWQPLVQOOLNLAQDLLEAVASKHN	433
QY	438	ETYSRRSRQAESTTDPMMEMTCGAGGEYSSENSITVAQVQYAYDNLRIINNILEDLS	497
DB	434	VTKPARGRRQRAVSSILID-----DDVYTAESALLTQIQFADMLRSQINTVLEELS	487
QY	498	KAWCREQHRALVWNLSEKINPTSVMSIYNRPVSAKRIGDVISVNSCIVVDQTSVSLHK	557
DB	488	RAWCREQHRASLMMNLSEKINPTSVMSIYGRPVSAKRIGDVISVNSCIVVDQSQVPLHR	547
QY	558	SLRLLSASD-EKCFSPRPVTFKFMNDSTIYKQOLGVNNEILLTTTYLETQENTYFYQA	616
DB	548	NNRVGPRDHAHEKSRPVPFINDSHLYKQOLGVNNEILLTTTALREVCHENTHEYFQG	607
QY	617	KTDVIYKNEYHLKTVPLSSITLDTFFIALNFTLLENVDKFIELYTRDEKRLSNVFDIE	676
DB	608	GNNMYFYKNYRVKTIPIPSAVATLDTFVNLNLTVENIDFQVIELYSREKRMSTVFDIE	667
QY	677	TMFREYNYAQRVSGRLKDLDLSTNRNQFVDAFGSLMDLGAVGQTVVNAVSGVATLFS	736
DB	668	TMFREYNYTQRTVGLRLDLSIATNRNQFVDAFGTLMDDLGVGKTVVNAVSSLATLFS	727
QY	737	SIVTGFIFIKNPPFGMLMIIVIGVLAIFYLTTKTIYETAPIKMIYPEIDKUKEREG	796
DB	728	SIVTGLIIFIKNPPFGMLIFGLAAVIAVILLRRRAAFANPANPQMIYPDQITKQKQ	787
QY	797	KSEIAPISEEELERIVLAMIHQNSHMETKRPD-----KDSILTRAQNLMLRKESG	849
DB	788	ENNVFPISEKHELDRLIMLAMDHDTQKDPDEKEGPESGGSANKANWLNAKLVLRRAAG	847
QY	850	YSLNKNAESVE 860	
DB	848	YQPLKRSOSTE 858	
RESULT 10			
O36362			
ID O36362; PRELIMINARY; PRT; 854 AA.			
AC O36362;			
DT 01-JAN-1998 (TrEMBLrel. 05, Created)			
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)			
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)			
DE Glycoprotein B.			
OS Alcelaphine herpesvirus 1 (wildbeest herpesvirus).			

OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Gammapesvirinae; Rhadinovirus.

OX NCBI_TaxID=35252;

RN [1] SEQUENCE FROM N.A.

RP STRAIN=CS00;

RC MEDLINE=97404659; PubMed=9261371;

RX Ensler A., Pflanz R., Fleckenstein B.;

RT "Primary structure of the alcelaphine herpesvirus 1 genome.;"

RL J. Virol. 71:6517-6525 (1997).

DR EMBL; AF005370; AAC58059.1; -

DR InterPro; IPR000234; Glycoprot. B.

DR Pfam; PF00606; Glycoprotein B; 1.

DR ProDom; PD00693; Glycoprot B; 1.

SQ SEQUENCE 854 AA; 97059 MW; 480D7CF352D5B59C CRC64;

Query Match 54.8%; Score 2489.5; DB 12; Length 854;

Best Local Similarity 57.0%; Pred. No. 3.1e-148;

Matches 471; Conservative 135; Mismatches 173; Indels 47; Gaps 8;

QY 64 NNRDSEON---KNYGSPTPEYRVCSAGVGDFRQTDHVCPCDASDMVHSEGLI 120

DB 42 SDNIIRQNRNTAKGIHSDPSAPFRVCSASNIGDIFRQTHSCPCNKDKEHNEGILLI 101

QY 121 YKQNIIPFMRVRYKRVKVVTTSTVYNGIYSDITNOHTFYKSIIBPWETEKMDTIYQCFNS 180

DB 102 FKENIVPVYKRVKRVKIVTTSTIYNGIYADAVTNQHVFSKSVPIYETRRMDTIYQCYN 161

QY 181 LRLNTGNNLLTYVDRDDINMTVFLQPDGVTPDVKRYGQPELYLEPGWFGSVYRRRTTV 240

DB 162 LDVTVGNNLLVYTDNGSNMTVDLQPDGLSNSVRRYHSQPEIHAEPGMLLGGYRRRTTV 221

QY 241 NCELMDMFARSPDPDFVTTATGTVMSPEW-----SGEDDHENKMKHEKVPVSV 291

DB 222 NCEVTDARAVPFRFYITNIGTIEMSPFWSKAMNETEESGPDRT-----LTV 272

QY 292 INNYKVVDYQNRGTVPGLKTRIFLDRBEYTLSEKHLKMSYCYPLTWKAFYNGIQTEHS 351

DB 273 AKDYRVVDYKFGTQPOGHTFIVDKBEYTLSSAQQFENISYCRWAHWSFDNAIKTEHG 332

QY 352 GSYHFVANDITASPTTSKEDMKENFTTYHCLNEEIKAEIEKKYAKVNSTHSGYGLDYKFK 411

DB 333 KSLHFVANDITASPTTNTQTRVLGKHVCLNLTIESELKSLRAKAVNDTHSPNGTAQYYL 392

QY 412 TDGGLYLVWQPLQNLNLDLDAKNKL-----NETYSRRSRQAESTTDPMMEMTCNG 462

DB 393 TNGGLLVWQPLVQOOLNLDLDAKGLLDVAKKQNTTTTTTTTSTRSRQRSSVS-----SG 444

QY 463 AGGEYSSENSITVAQVQYAYDNLRIINNILEDLSKAWCREQHRALVWNLSEKINPTSV 522

DB 445 IDVYTAESTILLTQIQFADVDTLRAQINNVLLELSRAWCREQHRASLMMNLSEKINPTSV 504

QY 523 MSMTYNRPVSAKRIGDVISVNSCIVVDQTSVSLHKSRLLSA-SDEKCFSPRPVTFKPMN 581

DB 505 MSSYGRPVSAKRIGDVISVNSCHVVDQDSVSLHSMRVPGDKTHECYSPRPVTFKFIN 564

QY 582 DSTIYKQOLGVNNEILLTTTYLETQENTYFYQAKTDMYIKMYEHLKTVPLSSITLTD 641

DB 565 DSHLYKQOLGVNNEILLTTTAVEICHENTHEYFOGNNMVFYKNYRVKTVMPVGDVATLD 624

QY 642 TFIALNFTLLENVDKFIELYTRDEKRLSNVFDIETMFREYNYAQRVSGRLKDLDLST 701

DB 625 TFWVLNLTVENIDFQVIELYSREKRMSTAFDIETMFREYNYTQRTVGLRLDLTLAT 684

QY 702 NRNQFVDAFGSLMDLGAVGQTVVNAVSGVATLFPSSIVTGTFINFKNPPFGMLMIIVIG 761

DB 685 NRNQFVDAFGSLMDLGLVGKTVLNAVSSVATLFPSSIVSGIIFIKNPPFGMLLFLGLIA 744

QY 762 VLFAIYELTKTIYETAPIKMIYPEIDKUKEREGKSEIAPISEEELERIVLAMIHQON 821

DB 745 WVITVILLNRKAKFAQNPQMIYPDITKITSREELQVDPISKHELDRLIMLAMDHYHAS 804

QY 822 SHMETK-----TRKPKDSILTRAQNLMLRKRSYSLNKNAESVE 860

Qy	673	FDIETMFREYNYAQRVSGLRKOLD--LSTNRNQFVDAFGSLMDDLGAVGQTVVNAVSGV	731
Db	629	LDIESMFREYNYTQRLSGISKELHNTVENNRDAIIRQFGNILDGLGNIGSVVNVASGV	688
Qy	732	ATLFSSIVTGFINFIPKPNFGMLMIIVIGVLFAIYFLTKTKIYETAPIKMIYPEIDKL	791
Db	689	FTLFGSIVTGLINFIKHPIGIPGIMTILIVGAVVLVFLNRNRMMYEAPIKMLYPHIDQS	748
Qy	792	KEREGKSEIAPISSEELERIVLAMHIHQONSH---METK--TRKDPKDSILTRAQNMLRK	846
Db	749	AQ---DDEVQIDQQQKLSILLAMHNFQBEHAKLIBEKSNNKXSLGSSLANIATKFLRN	805
Qy	847	RSGYSNLKNA 856	
Db	806	RKGYSALKTS 815	

RESULT 12			
Q66613	ID	Q66613	PRELIMINARY; PRT; 874 AA.
AC	Q66613		
DT	01-NOV-1996	(TEMBLrel. 01, Created)	
DT	01-NOV-1996	(TEMBLrel. 01, Last sequence update)	
DT	01-OCT-2001	(TEMBLrel. 18, Last annotation update)	
DE	Glycoprotein B.		
OS	Equine herpesvirus type 2 (strain 86/87) (EHV-2).		
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;		
OC	Gammaherpesvirinae.		
OX	NCBI_TaxID=82831;		
RP	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=95302501; PubMed=7783207;		
RA	Telford B.A., Watson M.S., Aird H.C., Perry J., Davison A.J.;		
RT	"The DNA sequence of equine herpesvirus 2.";		
RL	J. Mol. Biol. 249:520-528 (1995).		
RL	[2]		
RP	SEQUENCE FROM N.A.		
RA	Telford B.A.R.;		
RL	Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; U20824; AAC13795.1; --		
DR	InterPro; IPR000234; Glycoprot.B.		
DR	Pfam; PF00606; Glycoprotein B; 1.		
DR	ProDom; PD000693; Glycoprot.B; 1.		
SQ	SEQUENCE 874 AA; 98486 MW; DB03731D3B966410 CRC64;		

Query Match 43.0%; Score 1953.5; DB 12; Length 874;			
Best Local Similarity 46.1%; Pred. No. 1.7e-114;			
Matches 403; Conservative 147; Mismatches 250; Indels 75; Gaps 18;			

Qy	11	VLALWLYQVALYSLSTAETGVTSPTNTATWSTESPLTGHYTHDSSHGSGNNENRDS	70
Db	10	VLCMCV--AALCQGAQAEVWA-----ETTP----FATH-----RPEVVAA	46
Qy	71	EQNKNYIGSPSTPPYVCSAGV-GDVFRFQTDHVCPSDASDMVHSEGIILYKQNIIPFM	129
Db	47	ENPAPF-----UPFRVCAGSPTGGRIFRPLESCENTEDKHIEGIALYIKTNIPYV	101
Qy	130	FRYKRYKVVTTSTVNGIYSDSITNQHTFVKSGIEPMETEMDTIYQCFNSLRLLNTG	189
Db	102	FNVRKYRKIMTSTTIYKGVSEDAITNQHTRSAYVPLYEVQMDHYQCFSAVQVNEGHV	161
Qy	190	LTVYDRDDINMTVLPQVDCVTPDVKRYGSOPELYLEPGHFWGSGYRRRTTVNC	249
Db	162	NTYVYDRDGNNEATFLPADGLTSITRYQSQPEVYATPRNLLWSYTRTTTNC	221
Qy	250	RSNPFPDFVTAAGDVNESPFGSGEDDHENKWKHPWFVSVINNYKVVDYQNRGT	309
Db	222	RSMKPEFFVTSVGDITIEKSPFLKENGTEBEKILKRPHSIQLLKNYATKY----	277
Qy	310	K----TRIFLDREYTLSEKHLKNNMSYCLTLWKAFCYNGIQTEHSGSYHFVANDIT	365
Db	278	QADNATRFPAIFGDSLKWKATTENSQCDLIILWKGFNSAIQTOHNSSLHFANDIT	337


```
RESULT 14
Q9WRT9 ID Q9WRT9 PRELIMINARY; PRT; 829 AA.
AC Q9WRT9;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Glycoprotein B.
OS Macaca mulatta rhadinovirus 17577.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Rhadinovirus.
OX NCBI_TaxID=83534;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99174001; PubMed=10074154;
RA Searles R.P., Berguam E.P., Axthelm M.K., Wong S.W.;
RT Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF083501; AAD21335.1; --
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD00693; Glycoprot_B; 1.
SQ SEQUENCE 829 AA; 93300 MW; C21F406A5A194BAA CRC64;

Query Match 42.4%; Score 1927.5; DB 12; Length 829;
Best Local Similarity 44.2%; Pred. No. 7e-113;
Matches 388; Conservative 146; Mismatches 258; Indels 81; Gaps 14;

QY 11 VLALWLYQVALYSLSAETGVTSPNTATWSTESPLTGHYTHDSHGRGNENRDSE 70
DB :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 10 LLRAWV--IIAIGTAVGE-NVTTPKGATTAKTP-----42
QY 71 EQNKNIYGSPT-----PPYRVCASGVGDVFRFQTDHVCPCDASDMVHSEGIL 118
DB :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 43 -----GPSTPTPENPRAEAFKFRVCASASATGELFRNLEKTCPGTEKTHQEGIL 94
QY 119 LIYKONIIPFMRVRKRVKVTSTVYNGIYSDSITNQHTFYKSIETWETKMDTIYQCF 178
DB :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 95 MVFKKNIVPHIFKVRIRKRVKATSVTVYRGWTTAVTGKQEVIRPVQYIEINHMDTIYQCF 154
QY 179 NSRLNTGGNLLTVYDRDDINMTVFLQPVGDVTPDKRVYGSQBELYLEPGWFGWSYRRRT 238
DB :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 155 SSNRVNVNGIVNTYTRDFTNQTVFLQPVGELTDNIQRYFSQPVLYTTPGWFGIYRVRT 214
QY 239 TVNCELMDFARSNPPDFVATGDTVEMSPFWSG-----DDHENKMKHEKPFVSVI 292
DB :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 215 TVNCEIVDMIARSAEPYSYFVTALGDTVEVSPFCHNDSTCSVAEKTENGLGAR-----VL 269
QY 293 NNYKVVQYQNGTVPLCKTRIFLDREBYTLISWEKHLKMSYCPILTLKAFYNGIQIETHSG 352
DB :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 270 TNYTIVDFATR--OPTETRVFADSGEYTVSWKAEKDPKSAVCALTLLWKTIPRAIQTTHA 327
QY 353 SYHFVANDITASFTTSKEDKKEFNVTYHCLNEEIKAEIKYAKVNSHSGYDGLKYFKT 412
DB :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 328 SYHFVANDVATFSPISQVNTFNGTYPCLNDVLOKTLNATIKLSLTHATNSQYIET 387
QY 413 DGGLYLVWQPLIQRLLDANKLNKNETYSRRSRQAESTTDPMMWMTGNGAGGEYSSENS 472
DB :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 383 EGGFLWQPLTSLADREMLNGTT-----PAPPTTTANRVRRSGVINEQATDD 440
QY 473 ITVAOVAYDNRIRIRINNILEDLSKAWCRQEHRAALVNNELSKINPYSMTIYNRPVS 532
DB :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
DB 441 LAAPQLOFAYDKLRASINKVLEELSRACREQVDRDTYMWYELSKINPYSMTIYGRPVS 500
QY 533 AKRIGUVISVNCIIVDQTSVSLHLSLRSASDEKCFSPRPVTFKPMNDSTIYKGLGV 592
```

Qy	293	NNYKVVDYQNRGTVPGLGKTRIFLDREBYTLGWKHLKMSYCPFLTWKAFYNGIQTEHSG	350
Db	270	TNYTMDVATR--APTTETRFVADSGEYTSVWKAEDPKSAVCALLTWKTFPRAIQTHEEA	327
Qy	353	SYHFVANDITASFTTSKEDMKFEFNTYHCLNEEIKAEIEKKYAKVNSHYSKVGDLKYKPT	412
Db	328	SYHFVANDVATFTSPSEVANFTYSCLEDEVIQKTLNDTITKKLSDTHVTNGSAQYKKT	387
Qy	413	DGGLYLVMQPIQLRLLDAKNLNNETYS-----RRSRQAESTTDPMMEMTGN	461
Db	388	EGGLFLLWQPLTPLSLVDEMGLNGTTPAPPATTSTVSRRVRSVYNTNQATD-----	439
Qy	462	GAGGEYSSENSITVAQOYAVDNLRIPIRNNILEDSKAWCQBROHRAALVWNELSKINPTS	521
Db	440	-----NLAAPQLQFAYDKLRASINKLELSRAWCREQVRDTYMWYELSINPTS	489
Qy	522	VMSMIYNRPVSAGRIDGIVISYNSCIVDPOTSLSLHKSURLLSASDEKCFSPRPVTFKFWN	581
Db	490	WMTAIYGRPVSAKPVGDIAISVTDCAVDAQASVSIHKSUR--TSTPGMCIYSPRPVTFREL	547
Qy	582	DSTYKQGLGVNNBEILLTTTLYLETQCENTEYFYQAKTDMYIKYKNEYHLKTVPLSLSTILD	641
Db	548	STTLFGKGLGRNBEILLTDNQVEACKETCEHYFIASNTVYKYDVYFVKKINTSEISLIG	607
Qy	642	TFIALLNFTLLENVDFKVELYTRREKRLS--NVFDIETMFREYNYIAQRVSLGRLKDL----	697
Db	608	TFIALLNUSFIENIDFRVIELYSRAEKKLSGVSFDIETMFREYNYITQRLAGLRDLNNTI	667
Qy	698	DLSTNRNQFVDAFGSLMDLCAVGOTVNVNAVSGVATLFSSITVGTFINIKPFGGMLMI	757
Db	668	DL--NRDLARDLSIEVADLGDVGTVVNVASSVITLFGSIVSGFINIKPFGGMLMIL	725
Qy	758	VVIQVLFAIYELTKTKTIYETAPIKMIYPEIDKKKEREGKGEIAPISBEELERIVLAMHI	817
Db	726	VIVAVLIVFALNRRTAIAQAPIRMIYPDIDKMPGSGKK-----VDQEIKNVLAGMHQ	780
Qy	818	HQONSHMETKTRDKPKOSILTRAQNMRLKR--SGYSNLNKNASVE	860
Db	781	LQDEERRLDEQQRSAPSLFRASDGLKRRFRGKYLENEEAQE	824
RESULT 16			
Q9J2M4			
ID	Q9J2M4	PRELIMINARY;	PRT; 829 AA.
AC	Q9J2M4		
DT	01-OCT-2000	(TrEMBLrel. 15, Created)	
DT	01-OCT-2000	(TrEMBLrel. 15, Last sequence update)	
DT	01-DEC-2001	(TrEMBLrel. 19, Last annotation update)	
DE	ORF08.		
OS	Macaca mulatta rhadinovirus 26-95.		
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;		
OC	Gammaherpesvirinae; Rhadinovirus.		
OX	NCBI_TaxID=119193;		
RP	[1]		
RN	SEQUENCE FROM N.A.		
RC	STRAIN=Macaca mulatta rhadinovirus isolate 26-95;		
RX	MEDLINE=20173730; PubMed=10708456;		
RA	Alexander L., Denekamp L., Knapp A., Auerbach M.R., Damanian B.,		
RA	Desrosiers R.C.;		
RT	"The primary sequence of rhesus monkey rhadinovirus isolate 26-95:		
RT	sequence similarities to Kaposi's sarcoma-associated herpesvirus and		
RL	rhesus monkey rhadinovirus isolate 17577. ";		
RL	J. Virol. 74:3388-3398(2000).		
DR	EMBL; AF210726; AAF59985.1; -		
DR	InterPro; IPR00234; Glycoprot.B.		
DR	Pfam; PF00606; Glycoprotein B; 1.		
DR	ProDom; PD00693; Glycoprot B; 1.		
SO	SEQUENCE 829 AA; 93290 MW; 1459F972DB8B9EP4 CRC64;		
Query Match 42.3%; Score 1925.5; DB 12; Length 829;			
Best Local Similarity 44.2%; Pred. No. 9.3e-113;			
Matches 391; Conservative 144; Mismatches 246; Indels 103; Gaps 15;			

RESULT 17

Q9YTQ5

PRELIMINARY; PRT; 807 AA.

(TREMBLrel. 10, Created)

(TREMBLrel. 10, Last sequence update)

(TREMBLrel. 19, Last annotation update)

n B.

pesvirus 3.

DNA viruses, no RNA stage; Herpesviridae;

virinae; Rhadinovirus.


```

QY 604 ETCQNTYVYFOAKTDMYIKYKNEHLKTVPLSSITLTTFFIALNFTLLENVDFKVIETY 663
Db 577 EVCQHSCEHYFOGNGHIYKYKNEYVSTWNLTDVPLTMTLNLSLVENVDFQVQLVS 636
QY 664 RDEKLSNVDFDIETMFREYNVYVYQVSLRKDLDD-LSTNRNQFVDAGSLMDLGAQGO 722
Db 637 QKEKLSNVDFDIETMFREYNVYVYQVSLRKDLDDSIHGRDSFQFLGDLVQDLVPPVD 696
QY 723 TVNVASGVATLFSSTVITGFINFIKNPFGGMLMIWIGVLFALY-FLTKTKIYETAPI 781
Db 697 VIVNVASGVSLFSGIVSGVISFLKNPLGAILITIALIVGGIIVLYLFIETRSTVYQ-API 755
QY 782 KMIYPEIDKLKREGKSEIAPISEBELERIVLAMIHQNSHMETKTRDKPK-----D 834
Db 756 RMLYPEVD---RAPQNVQPIPEQVRSILLAMEHFOQOQOQOQOQOQOQOQOQOQO 811
QY 835 SILTRAQNMRLKX---SGSYNLKNAES 858
Db 812 TIRESTSNILRRRGGGYTRLRQDS 838

RESULT 19
Q93K7
ID Q93K7 PRELIMINARY; PRT; 870 AA.
AC Q93K7;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE ORF4
OS Calitricrine herpesvirus 3.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammapherpesvirinae; Lymphocryptovirus.
OX NCBI_TaxID=106331;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CJ0149;
RX MEDLINE=21107697; PubMed=11158621;
RA Cho Y., Ramer J., Rivailier P., Quink C., Garber R.L., Beier D.R.,
RA "An Epstein-Barr-related herpesvirus from marmoset lymphomas.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:1224-1229(2001).
DR EMBL; AF319782; AAK38211.1; -
DR Interprot; IPR000234; Glycoprot B.
DR Pfam; PF00606; Glycoprotein B; 1.
DR ProDom; PD000693; Glycoprot B; 1.
SQ SEQUENCE 870 AA; 98634 MW; 6B7F0E0A7A0E8450 CRC64;

Query Match 40.5%; Score 1841; DB 12; Length 870;
Best Local Similarity 44.3%; Pred. No. 2.1e-107;
Matches 371; Conservative 164; Mismatches 233; Indels 70; Gaps 17;

QY 80 PSTFYRVCASGVGDVRFQTDHVCBPAS-DMVHSEGILLIYKQNIIPFMFRVKYRKV 138
Db 45 PYEFPFRVCEMNHGDIIRFATRISCPNFNLQENHTEGLIVFKDNIIPIYFNVRAYTKV 104
QY 139 VTTSTVNGIYSDSITNHTFYKSTPWTETKMDTIYOCENSLRLNTGNTLLTYVDRDDI 198
Db 105 VTTVTVNGWYADAVTNAHEERYSPDYLENNMDTYOCYNSVRMVKAGVERVYVDRDK 164
QY 199 NMTVFLQPDVGPVDPVKRYGSOPELYLEPGFWGFSYRRRTTVCNCLMDMFARSNPPDF 258
Db 165 NITVNLRPASGISLTVRRVNSQPDLYDDPGWVWYRRTTVCNITDMQAKSNSPFF 224
QY 259 VTATGDTVMSPFSGEDDHENKMEKWPVSVINNYKVQYQNGTVPLGKTRIFLDRE 318
Db 225 VTTLGQTVMSFPFYNG--SNSETFNEPAGTFEIREKYIVTFENRGTIPKBEKRAFLNK 282
QY 319 EYTLSEKHLKNWSYCPILTAKAFYNGIOTESHGSHYFVANDITASFTTSKEDMKFN 378
Db 283 TTYLSKULDARS-SYCLPTMMKDYSAIRTDGSGSFHPTPGQTASFVT-----NT 333
QY 379 -----YHCLNEIKAEIKKYAKVNSTHSA-YGDLKYFKTDGGLVYLWQPLIQNRLD 430

```

```

Db 334 HEKLPDOFKCIEQVKNKTIIEQFAKLSKNKYDRDAEIQYFRTAGLLIAMLVPTPKSLIT 393
QY 431 AKNKLNNETYS-----RBSROAES-----TTPDMEMTGNG 462
Db 394 IRDIALNNTAEQDGAARDQPDPTANSPTPOAPPLRRNRREAGSDRNTNTITDVSATTEG 453
QY 463 AGGEYSSENSI---TVAQVQVAYDNLRIRINNILEDLSKAWCREQHRALVYNLSKINP 519
Db 454 TGDKDSILGTIDNPATAIQIOPAYDSLSRSQINRMGLDGLASAWCQQRQNMVNLTKINP 513
QY 520 TSVSMIYNRPVSAKRTGDVIVSNCIVVDQTSVLSLHSLRLLSASDEKCSRPVPTKPF 579
Db 514 TTVMTGVYGVPAAKRGIDVISVSRCPVEQESVLSRSMRVL-GSETSCYTRPLVSFSF 572
QY 580 MNDSTIYKGLQGVNNEILLTTTLYLETQENTYFYFOAKTDMYIKYKNEHLKTVPLSSITT 639
Db 573 PNDSTYEGQLGQNEILLTKMVENCQDTQCHTFQSGNEMHVRFDVQHFXTIIPVEDVAT 632
QY 640 LDTFIALNFTLLENVDFKVIETYTRDEKRLSNVFDIETMFREYNVYVYQVSLRKDLDD- 698
Db 633 LOTFITLNTSFITNIDFQTLLEYSDERRASNVFDIEGIFREYNFQIQNIANLRDLDA 692
QY 699 LSTNRNQFVDAGSLMDLGAQVQTVNVASGVATLFSSTVITGFINFIKNPFGGMLMIIV 758
Db 693 ISNNRQFVDGLGELTDSLGAVGQAITNIVSSVGLFSLISGFSIFPKNPFPGVLIILAG 752
QY 759 VIGVLFALYFLTKTKIYETAPIKMIYPEIDKL-KERE-GKSEIAPISEBELERIVLAMI 816
Db 753 IVAVIFLVIFLVQRNRTIAQOPVQFLYPEIQSLTKNREIQGSNVAPIKQQLDAIMLALY 812
QY 817 IH-QQNSHMETKTRDKPKDSILTTRA-----QNNLKRSGYSN-----LKNAESVE 860*
Db 813 EQTAKRTQAESKT-SDSTTSLPSRALEARNRLRLKPKGRENRSDLKPLLTNIEDTE 869

RESULT 20
Q8UZD5
ID Q8UZD5 PRELIMINARY; PRT; 864 AA.
AC Q8UZD5;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE BALF4.
OS Cercopithecine herpesvirus 15.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammapherpesvirinae; Lymphocryptovirus.
OX NCBI_TaxID=104228;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LCL8664;
RX MEDLINE=97048062; PubMed=8892903;
RA Franken M., Devergne O., Rosenzweig M., Annis B., Kieff E., Wang F.;
RT "Comparative analysis identifies conserved tumor necrosis factor
RT receptor-associated factor 3 binding sites in the human and simian
RT Epstein-Barr virus oncogene LMP1.";
RL J. Virol. 70:7819-7826(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=LCL8664;
RX MEDLINE=99412410; PubMed=10482645;
RA Rivailier P., Quink C., Wang F.;
RT "Strong selective pressure for evolution of an Epstein-Barr virus
RT LMP2B homologue in the rhesus lymphocryptovirus.";
RL J. Virol. 73:8867-8872(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=LCL8664;
RX MEDLINE=20304984; PubMed=10846073;
RA Jiang H., Cho Y.G., Wang F.;
RT "Structural, functional, and genetic comparisons of Epstein-Barr virus
RT nuclear antigen 3A, 3B, and 3C homologues encoded by the rhesus
RT lymphocryptovirus.";

```


Qy	687	QRVSGLRKDLID-LSTNRNPFVDFAGSMDDLGAVGQTVVNAVSGVATLFSIVTGFINF	745		
Db	668	YSLAGIKKIDLNTIDYNRDLVQDLSDMADLIGRNVNVSSVVTFFSSIVTGFIFK	727		
Qy	746	IKNPFQGLMLIIVVIGLVFAIFYFLTKTKIYETAPIKMIYPEIDKLKEREG---KSEIFA	801		
Db	728	FTNPLGGIFILLIIGGIIFLVVNLNRRNSQPHDAPIKMLYPSVENYAAQAQPPVSA	787		
Qy	802	PISEELERIVLAWH-IHQNSHMETK-TRKDPKDSILTIAQNMLR-KRSGYSNL	853		
Db	788	AIDKEEIKRILLGMHQVQBEKEAQKQLTNSGP--TLWQKATGFLNRNRKGYSQL	840		
RESULT 22					
Q65561	PRELIMINARY; PRT; 874 AA.				
AC	Q65561;				
DT	01-NOV-1996 (T-EMBLrel. 01, Created)				
DT	01-NOV-1996 (T-EMBLrel. 01, Last sequence update)				
DT	01-OCT-2002 (T-EMBLrel. 22, Last annotation update)				
DE	Glycoprotein B.				
GB	GB.				
OS	Bovine herpesvirus 4.				
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;				
OC	Gammaherpesvirinae.				
OX	NCBI_TaxID=10385;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=66-P-347;				
RX	MEDLINE=951176575; PubMed=7871762;				
RA	Goltz M.M., Broll H.H., Mankertz A., Weigelt W.W., Ludwig H.H.,				
RA	Buhk H.H., Borchers K.K.;				
RT	"Glycoprotein B of bovine herpesvirus type 4: its phylogenetic				
RT	relationship to GB equivalents of the herpesviruses.";				
RL	Virus Genes 9:53-59(1994).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=20583805; PubMed=11152491;				
RA	Zimmermann W., Broll H., Ehlers B., Buhk H.-J., Rosenthal A.,				
RA	Goltz M.;				
RT	"Genome Sequence of Bovine Herpesvirus 4, a Bovine Rhadinovirus, and				
RT	Identification of an Origin of DNA Replication.";				
RT	J. Virol. 75:1186-1194(2001).				
DR	EMBL; Z15044; CAA78761.1; --				
DR	EMBL; AF318573; AAK07927.1; --				
DR	InterPro; IPR000234; Glycoprot B.				
DR	Pfam; PF00606; Glycoprotein_B; 1.				
DR	ProDom; PD000693; Glycoprot_B; 1.				
KW	Membrane.				
SQ	SEQUENCE 874 AA; 99419 MW; 95CCBA04EB7F1360 CRC64;				
Query Match 39.3%; Score 1787; DB 12; Length 874;					
Best Local Similarity 43.5%; Pred. No. 5.2e-104;					
Matches 367; Conservative 159; Mismatches 248; Indels 70; Gaps 20;					
Qy	34	SPDNTATWSTESPLTGHYCHTHDSHGERGNENRDSEEQNKNIYGSSTPPRYVCSA-SG	92		
Db	58	SPONT---STSKPSTNQGTSTFT-----IPTVTDTSKKNFY-----KYRVCSSASS	102		
Qy	93	VGVDFRFQTDHVCPDASDMVHSEGIILYKQNIIPFMFRVKRYKVVTTSTVYNGIYSDS	152		
Db	103	SGELFFRFDLDQICPDYTKDKKHGVEGILLVLKKNIVPIFKYRKYKIATSVTVYRGWQA	162		
Qy	153	ITNQHTFYKSIPEWTEKMDTIYQCFNSRLNTGGNLLTYVDRDDINMTVLPQVDGVTP	212		
Db	163	VFNRRDDISRAIPYNEISMTIDRTYCHPSAMATVINGILNTVIDRSENKSVLPQVAGL	222		
Qy	213	DVKRYGSPLELYPCGFWGYSYRRRTVNCCLMDMFARSNPPDPFFVTATGDTIVEMSPFW	272		
Db	223	NNRVFSQPLIIAEPGWGFFCIVRVRTVNVCEVVDYMARSEVPTHTFTALGDITIEISPPC	282		
Qy	273	SGEDDHENKM-----HEKPFVFSVINNYKVVDYQNRGTVPVLPCKTRIFLDREEVTL	322		

RT "Cloning and characterization of rhesus cytomegalovirus glycoprotein
B.",
RL J. Gen. Virol. 78:2009-2013(1997).
DR EMBL; U59238; AAB70023.1; - - - - -
DR InterPro; IPR000234; Glycoprot_B.
DR InterPro; IPR000847; HTH_LysR.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
DR PROSITE; PS00044; HTH_LYSR_FAMILY; 1.
SQ SEQUENCE 854 AA; 97859 MW; 7AAEACAEADD9F69D1 CRC64;

Query Match 24.8%; Score 1126; DB 12; Length 854;
Best Local Similarity 31.5%; Pred. No. 26-62;
Matches 281; Conservative 172; Mismatches 336; Indels 102; Gaps 27;

QY 20 VALYSLSTAETGVTSS---PPNTATWSTESPLTGHGTHDSHGCGRNNENRDSERQNKI 76
DB 15 VVYVSIASSSTGTASGAVTSPTEITGGLI-----ENTTLRTHVEFKINM 61
QY 77 YGSPSTFPYRVCSAGVGDFVFQTDHVC---PDASDMVHSEGLILLYKONIIPFMRV 132
DB 62 ----SKFPYRVCSMAQGTDLRFEQINCDSEKPKEDF--EEGMVVKYKDIRPYTKV 115
QY 133 RYKRVVTTSTVYNGIYSDSINTQHTFYKSIETETKMDTIYQCFNSLRNLNTGGLTY 192
DB 116 HMYQKILTFQSYSTIRENHLGLFSQHLAVPMWEVHYINKLRNCYNSVVRNVAG--ATY 173
QY 193 VD--RDD--INMTVFLOPQDGVTPDKRVYGSQBELYLEPGWFGVSRRTTNCELMDMFA 249
DB 174 VNYHRDSYVNEFMLLVEDDFSNTHSSRFVTVKELHFKGSTW--LYTTSNCVNCMVTVTTA 232
QY 250 RSNPFPDFVTATGTVMSPWSEDD--HENKMEKPFVSVINNYKVVDYQNRGTVP- 307
DB 233 RSKYDYDFVTSQKVDVISPYNSSNKHFGENDK---FSVRKNYSMIAYGRDPAE 289
QY 308 ----LGKTRIFLDREYTLSEKHLKNMSYCPLTLKAFYNGIOTERHSGSYHFVANDIT 362
DB 290 VAHPLVG---PFRPDVLMWDIVVEAANTCEYTFWEQSERTIRSEADDIHYHTSSMT 345
QY 363 ASFTTSKEDMKFNTYHCL-----NEETKAEIKKYAKVNSTHSGYDGLKPKTDGGLYL 418
DB 346 ATFLTSKELNESDPFCIKDKANEOLQLFNTSY---NETYQSGVSNVYETGGIIV 402
QY 419 VMOPLIONRLDKN---KLNNETYSRRSRQAESTTDPMMEMTNGAGGEVSSP--NSI 473
DB 403 FWPVKEKSIEMELAVAYNNTSSTRKSTDSASD-----SNKITEEVLKSI 452
QY 474 TVAQVQYAYDNLRIINNILEDKAWCREQHRALVWNELSKINPTSVMSIYNRPVSA 533
DB 453 VYAOQYIYDYLRLNINRALQIAEAWCKDQKRTLEVEFKELSKINPSAMLSAIYDKPIAA 512
QY 534 KRIGDIVSNICIVDQTSVLHLSLRLLSASDEKCFRPPVTFKPMNDSTIYKQGLVN 593
DB 513 RFGVDVSLAKCQVEVDQNSVKVLRDMR--TKESGVCYSPRVVLYTFKNSHVQYQGLGEY 570
QY 594 NEILLTTLTYLETQENTEYFOAKTDMYIKNYEHLKTVPLSSITLTLDTFALNFTLEN 653
DB 571 NEILLGRHTEACEYPSKIYAGNSSEYVDLYKRMIPLODSTVDTMISLSDIDPLEN 630
QY 654 VDFKVIETYLRDEKLSNVFDIETMFREYNYAQRVSGLRKDLDLSTNRNQVDAFGSL 713
DB 631 TDFKALELYSEDELSSNVFLEDIMREFNTYKQRMHVEGKVPDKVPG---YLRGLDDM 687
QY 714 MDDLGAQGVTVNAVSGVATLSSIVTGFINFKNPPFG---MLMIIVIGVFAIYPLT 770
DB 688 MSGLSAGKALGAIVGAGGVASVEGVGFIKNPPFGSFVILPFLAVGLVILYMRQ 747
QY 771 KKT--KIVE-----TAPIKWIPEIDKLKREGKSEIAP-----ISEELERIVLAWH 816
DB 748 KRAYEKPEHPFPYVVPPTVKEAPPSEYQGYENIKEAAPATKFELEEAYQMLLA-- 805
QY 817 IHQNSHMETKTRDKPKD-----SILTRAQNLMLKRSYSLNKLNAESV 859

DB 806 LQKLDQEKRRKAEADDEDFAHQAGFLDLRLN--RRRGYQKIQNEYEV 854

RESULT 28
Q69169 PRELIMINARY; PRT; 906 AA.
ID Q69169
AC Q69169;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Glycoprotein B.
OS Human cytomegalovirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10359;
RN. [1]
RC SEQUENCE FROM N.A.
RP STRAIN=C325A;
RX MEDLINE=92230241; PubMed=1314465;
RA Chou S.;
RT "Comparative analysis of sequence variation in gp116 and gp55
components of glycoprotein B of human cytomegalovirus.";
RL Virology 188:388-390(1992).
DR EMBL; M60927; AAA45926.1; - - -
DR InterPro; IPR000234; Glycoprot_B.
DR InterPro; IPR000847; HTH_LysR.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
DR PROSITE; PS00044; HTH_LYSR_FAMILY; 1.
SQ SEQUENCE 906 AA; 101833 MW; 523C585377AAC9F2 CRC64;

Query Match 24.7%; Score 1123.5; DB 12; Length 906;
Best Local Similarity 30.9%; Pred. No. 3-le-62;
Matches 286; Conservative 182; Mismatches 353; Indels 105; Gaps 24;

QY 14 LWLYQVALYSLSTAETG--VTSPNPTATWSTESPLTGHYGTDS--HGBRG 62
DB 5 IWCL--VVCVNLICVLGAASVSSSTSHATST-----HNGSHTSRTTSACTRSVSSQHV 56
QY 63 NENRDEEONKNTYG-----SPSTFPYRVCSAGVGDFVFQTDHVC-----PD 107
DB 57 TSSEAVSHRANETIYNTLTLYGVDVGVNTTKYPRVCSMAQGTDLIRFERNIVCTSMKPI 116
QY 108 ASDMVHSEGLILLYKONIIPFMRVRYKRVKVTSTVYNGIYSDSINTQHTFYKSIETPE 167
DB 117 NEDL--DEGMVVKYKRNIVAHTFKRVYQKVLTFRRSYAIIHTYLLGNSNTEYVAPPWE 174
QY 168 TEKMDTIYQCFNSLRLNTGGLTYVDORDDI--NMTFELQPDVGVTPDKRVYGSQPELYLE 226
DB 175 IHHNSHSQCYSSYSRVIAGTVFVAYHRDSYENKTMQIMDDYNTSHSTRVTVKDWHS 234
QY 227 PGWFWGYSRRRTTVCNCELMDFARSNPPDFVTFATGDTVMSPWSEGEDDHENKMEKP 286
DB 235 RGSTW--LYRETCLNLCMVTTITARSKYPHFFATSTGDDVVDISPFGYNGTNRNASYFGENA 293
QY 287 WFSVINNYKVVDYQNRGTVPLGKTR--IFLDREYTLSE--KHLKNMSVCLTLKAFY 343
DB 294 DKFFIFPNYITVDFGRPNSALETHRSVAFLEKADSVISWDIODEKNT--CQLTFWEASE 352
QY 344 NGIQTESGSHYFVANDITASFTTSKEDMKFNTYHCLNEETKAEIKKY-AKVNGTHS 402
DB 353 RTIRSEADESYHFSACKMTATFLSKQEVNMSDALDCVDEAKINKLOQIFNTSYNTYE 412
QY 403 KYGDKLYFKTDGGLYLVWQPLIQNRLLDAKNLNNTYSRRSRQAESTTDPMMEMTNG 462
DB 413 KYGNVSVFETGGVLVFWQGIKQKSLVELERLANRSSLNLTNRKST-----NGNN 465
QY 463 AG--GEVSSNSITVAQVAYDNLRIINNILEDLSKAWCREQHRALVWNELSKINPT 520
DB 466 ATHLSNMESVHNLVYAOQYTYDTRLGYINRALAQIAEAWCVDRRSLEYFRELSKINPS 525
QY 521 SVMSIYNRPVSAKRIGDIVSNICIVDQTSVLSHLKSLRLLSASDEKCFRPPVTFKFM 580


```
DR InterPro; IPR000847; HTH_LySR.
DR Pfam; PF00606; Glycoprotein B; 1.
DR ProDom; PD000693; Glycoprot B; 1.
DR PROSITE; PS00044; HTH_LySR_FAMILY; 1.
SQ SEQUENCE 905 AA; 101916 MW; 90BEE74D7581EFF1 CRC64;

Query Match 24.6%; Score 1119; DB 12; Length 905;
Best Local Similarity 30.9%; Pred. No. 5.9e-62;
Matches 286; Conservative 186; Mismatches 349; Indels 104; Gaps 25;

QY 14 LWLYQVALYSLSIAETG--VTSPTNTATWSTESPLTGHYGHDS-----HGERGUNE 65
DB 5 IWCL--VVCNLCI VCLGAUVSSSTRAT-----STHNGNHTSHSTSAQTRSISSQVTS 58
QY 66 NRDSSEQNKNIYG-----SPSTPPYRVCSGASGVDFRFTDHC-----PDASD 110
DB 59 EAVSHRANETIYNTLKYGDDVGVVNTTKPYRVCSMAQGTDLIRDRNIVCTPMKPIED 118
QY 111 MVHSGILLIYKONIPFMRVRKRVKVVTSVTVNGIYSDSITNQHTFYKSIETPWEK 170
DB 119 L--DEGIMVYKRNIVAHTEKRVYQKVLTPRRSYAVIHTTYLLGNSNEYVAPPWMEI 176
QY 171 MDTIYOCENSLRNTGNNLLTVDRDDI--NMTVFLQPDGVTPDKRYGSGQPELYLEPGW 229
DB 177 INSHSQCYSSYSRVIAGTVFVAHYHRDSYENKTMQLMPDDYSNTHSTRYVTVKDDQ 236
QY 230 FWGYSRRRTTVMCELMDFARNSPPDFVTATGTVEMSPFWSGEDDHENKMKHPWFV 289
DB 237 TW-LYRETCNLNCMTIITARSKYPIYHFPATSTGVDVISPYPYNTNRRNASVFG 295
QY 290 SVINNYKVDVQNRGTPLGKTRI--FLDREBYTLSE--KHLKNMSYCLPLWKAFYNGI 346
DB 296 FIFPNYTVISDRGRNSALETHRLVAFLERADSVISWDIQDEKNVT--QLTFWEAS 354
QY 347 QTEHSGSVHFVANDITASFTTSKEDMKENITYHCLNEEIKAEIKKY--AKVNSTH 405
DB 355 RSEAESYHFSAKMTATFLSKQKVNMSSDALDCVRDEALNKLQOIFNASTYQTYEY 414
QY 406 DLKYEKTDGGLYLVNQPIQLRLLDAKNKLN---NETYSRRSRQAEISTDPMMETGN 461
DB 415 NVSVFETGGVVVFGQIKQKSLVLELRANRSSLNLTSH--RTRSTGTGNTVTHL-- 468
QY 462 GAGGYSSENSITVAQVAYDNLRIINILEDLSKAWCREQHRRAALVWNLSEKINPTS 521
DB 469 ---SNMDSVHNLVYAQLQFTYDTRGYNRALQTAEAWCVDQRTLEVFEKELSKIN 525
QY 522 VMSMYNRPVSAKRIGDVISVNCIVDQTSVLSHKSRLLSASDEKCFSPRPVTFKPM 581
DB 526 ILSAIYNKPIAARFNGDVLGASCVTINQTSYKVLDRMN--VKESPGRCYSRPVVF 584
QY 582 DSTIYKGLGVNNEILLTTTLETQENTYFYQAQKTDWYIKNYEHLKTVPLSSITLD 641
DB 585 SSYVOYGQGEDNELLGNHRTTEQLPSLKIFAGNSAYEYVDYLFKRMIDLSSIS 644
QY 642 TFIALNFTLLENVDFKVIETRYDEKRLSNVFDIETMFREYNYAQRVSGRLKDL 701
DB 645 SWIALDIDPLENTDFRVLLEYLSQKELRSSNVFDLEEIMREFNYSYQRYKVEDK 701
QY 702 NRNQVDAFGSLMDDLGAAGVQTVNNAVSGVATLFSISVTGFINFKNPFGLM 761
DB 702 PLPPYLKGLDMLMSGLGAAGKAVGVAIGAVGAVASVVEGVATFLKNPFGAFT 761
QY 762 VLFALYFLTKTKIYETAPIKWIYPEI-----DKLKE-----R 794
DB 762 VVIILYITQRRLCMQPLQNLFFYLVASDGTTVTSNGTKDTSLOQAPPSVEES 821
QY 795 EG-----KSEIAPISSEELERIVLAM-----HIHQNS-----HMETKTR-- 835
DB 822 KGPGLPPSDASTAAPPYTNEQAYQMLLALVRLDAEQRAQQGTDSDLCQTGTQ 881
QY 836 ILTRAQNMLRKRGVSNLKNAESVE 860
DB 882 LLDRLRH---RKNGYRHLKDSDEE 903
```



```

DT 01-DEC-2001 (TREMELrel. 19, Last annotation update)
DE Glycoprotein B.
OS Human herpesvirus 7.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae.
OX NCBI_TaxID=10372;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AL;
RX MEDLINE=974113886; PubMed=9268767;
RA Secchiero P., Berneman Z.N., Sun D., Nicholas J., Reitz M.S. Jr.;
RT "Identification of envelope glycoproteins H and B homologues of human
RT herpesvirus 7.";
RL Intervirology 40:22-32(1997).
DR EMBL; AF007829; AAB63200.1; -.
DR InterPro; IPR000234; Glycoprotein B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprotein_B; 1.
SQ SEQUENCE 822 AA; 93176 MW; 052F7B523AD257D1 CRC64;

Query Match      24.0%; Score 1090.5; DB 12; Length 822;
Best Local Similarity 33.7%; Pred. No. 3.2e-60;
Matches 273; Conservative 146; Mismatches 314; Indels 77; Gaps 24;

Qy 78 GSPSTFPRVCSASGVDFRQTDHVCDD-ASDMVHSEGLLIYKQNIIPFMRVKYR 136
Db 78 GSPSTFPRVCSASGVDFRQTDHVCDD-ASDMVHSEGLLIYKQNIIPFMRVKYR 136
Qy 28 GHNOHLPKRICSIATGDLVRPDRVSCASVSGNIKTTEGILIIYKTKIEAHTFSVRTFK 87
Db 28 GHNOHLPKRICSIATGDLVRPDRVSCASVSGNIKTTEGILIIYKTKIEAHTFSVRTFK 87
Qy 137 KVVITSTVYNGI-----YSDSITNQHTFYKSTPWEKMDYIYQCFNSLRNTG--GNLL 190
Db 137 KVVITSTVYNGI-----YSDSITNQHTFYKSTPWEKMDYIYQCFNSLRNTG--GNLL 190
Qy 88 KELTFQITTYRDVGTGYFLDRTVT--TLPMPIE--EVHVMNTEARCLSSIVKRSSEBEYV 143
Db 88 KELTFQITTYRDVGTGYFLDRTVT--TLPMPIE--EVHVMNTEARCLSSIVKRSSEBEYV 143
Qy 191 TVVDRDDNMVTFLOPVDGVTVDKRYGSOPELYLEPGWFGSXYRRRTTVNCELMDMEAR 250
Db 191 TVVDRDDNMVTFLOPVDGVTVDKRYGSOPELYLEPGWFGSXYRRRTTVNCELMDMEAR 250
Qy 144 AYHDEYVNTKLDLILPDKSDTVRYITTKEPFLUNGPLV-FYSTSINCIVTDCAK 202
Db 144 AYHDEYVNTKLDLILPDKSDTVRYITTKEPFLUNGPLV-FYSTSINCIVTDCAK 202
Qy 251 SNPPDFPVTATGDTVEMSPWMSGDDDE--NKMHEKPMFVSVINNY---KVDYQNRGTV 306
Db 251 SNPPDFPVTATGDTVEMSPWMSGDDDE--NKMHEKPMFVSVINNY---KVDYQNRGTV 306
Qy 203 TKYPDFPALTGTETVEGSPFNGINSKTFNEPTEKILF---RNNYTLKTFDDGSKGNF 259
Db 203 TKYPDFPALTGTETVEGSPFNGINSKTFNEPTEKILF---RNNYTLKTFDDGSKGNF 259
Qy 307 PLGKTRIFLDREBYTLSEKHLKMSYCPFLWKAFYNGIOTEGSGSVHFVANDITAGFT 366
Db 307 PLGKTRIFLDREBYTLSEKHLKMSYCPFLWKAFYNGIOTEGSGSVHFVANDITAGFT 366
Qy 260 VTLTKMAFLKGNITTFSEVQNEESSICLLKHMTWIPHALRAENANSHPHIAQELTASFV 319
Db 260 VTLTKMAFLKGNITTFSEVQNEESSICLLKHMTWIPHALRAENANSHPHIAQELTASFV 319
Qy 367 TSKEDEKFNNTYHCLNEEIKAEIEKKY-AKVNSTHSKYGDLKYPKTDGGLVYVWQPLIQ 425
Db 367 TSKEDEKFNNTYHCLNEEIKAEIEKKY-AKVNSTHSKYGDLKYPKTDGGLVYVWQPLIQ 425
Qy 320 TKSNYTLSDSKYNCINNYISILDEIYQTNNSHDKNKSGVEYFPTGGLILWQPLIQ 379
Db 320 TKSNYTLSDSKYNCINNYISILDEIYQTNNSHDKNKSGVEYFPTGGLILWQPLIQ 379
Qy 426 NRLLDAKNLNNTYSRRSRQAESTTDPMMEMTGNAGGYSSENSITVAQVQVAYDNL 485
Db 426 NRLLDAKNLNNTYSRRSRQAESTTDPMMEMTGNAGGYSSENSITVAQVQVAYDNL 485
Qy 380 RKLTVLEN-FSNASKRKRK-----ELETNKDIVVYVQLQVLYDTL 418
Db 380 RKLTVLEN-FSNASKRKRK-----ELETNKDIVVYVQLQVLYDTL 418
Qy 486 RIRINILEDLSKAWCRQHRALVWNLKINPTSVMSMIYNRPVSKRIGDVISVNC 545
Db 486 RIRINILEDLSKAWCRQHRALVWNLKINPTSVMSMIYNRPVSKRIGDVISVNC 545
Qy 419 KDYINTALGLAEAWCLDQKRTITVLHLSKISPSGIISAVYGPMSAKLIGDLVAVSKC 478
Db 419 KDYINTALGLAEAWCLDQKRTITVLHLSKISPSGIISAVYGPMSAKLIGDLVAVSKC 478
Qy 546 IVDQTSVSLHSLRLLSASDE---KCFSPRPVTFKMNDS-TLYKGLGVNNEILLTTT 601
Db 546 IVDQTSVSLHSLRLLSASDE---KCFSPRPVTFKMNDS-TLYKGLGVNNEILLTTT 601
Qy 479 IEVQTSVQLHSLKRLTKDSDYALRCYSRPLLTYSFANSKETYVLGGLDNEILLGNH 538
Db 479 IEVQTSVQLHSLKRLTKDSDYALRCYSRPLLTYSFANSKETYVLGGLDNEILLGNH 538
Qy 602 YLETG-QENTYIYFOAKTDMYIYKNEYHLKTVPLASSITTLDTFIALNFTLLENVDFKVE 660
Db 602 YLETG-QENTYIYFOAKTDMYIYKNEYHLKTVPLASSITTLDTFIALNFTLLENVDFKVE 660
Qy 539 RTECEQSNTKIFLSGKF-AHFKDYTVYVNSLITEIEALDAFVDNLIDPLENVDFTLLE 597
Db 539 RTECEQSNTKIFLSGKF-AHFKDYTVYVNSLITEIEALDAFVDNLIDPLENVDFTLLE 597
Qy 661 LYTRDEKLSNVFDIETMFRYNYAQRVGLRKDLDLSTNRNQVDFAGSLMDDLGAV 720
Db 661 LYTRDEKLSNVFDIETMFRYNYAQRVGLRKDLDLSTNRNQVDFAGSLMDDLGAV 720
Qy 598 LYTKDELSKANVFDLETILREYNSYK---SALHHIETKIATVPTIYIGGIDTFFKGLCAL 654
Db 598 LYTKDELSKANVFDLETILREYNSYK---SALHHIETKIATVPTIYIGGIDTFFKGLCAL 654
Qy 721 GOTVNAVSGV-ATLFSIVTGFNFIKPNPGGLMIIVIGLVFAIYFELTKYKIYETA 779
Db 721 GOTVNAVSGV-ATLFSIVTGFNFIKPNPGGLMIIVIGLVFAIYFELTKYKIYETA 779
Qy 655 G-LGLGAVGLVTAGALGVNVGVSFLKNPFGGALTILLTGLVIGLVFLRHKRLAQ 713
Db 655 G-LGLGAVGLVTAGALGVNVGVSFLKNPFGGALTILLTGLVIGLVFLRHKRLAQ 713
Qy 780 PIKMIYPEIDKLER--EGKSEIAPISBEELERIVLAMIHIHQONSHMETKTRKD--PKDS 835
Db 780 PIKMIYPEIDKLER--EGKSEIAPISBEELERIVLAMIHIHQONSHMETKTRKD--PKDS 835

Db 714 PIDILFPYTSKSTNSVLQATQSQVQAQVKEPLD-----SSPPYLKTKNKTDEPQGD 762
Qy 836 ILTRAQNMLRKRCYSNLSKNAESVEMLNTL 865
Db 763 DITHTNE-----YSQV---EALKMLKAI 782

RESULT 35
OSV615
ID Q8V615 PRELIMINARY; PRT; 859 AA.
AC Q8V615;
DT 01-MAR-2002 (TREMELrel. 20, Created)
DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMELrel. 21, Last annotation update)
DE Glycoprotein B.
GN GB.
OS Porcine cytomegalovirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=109993;
RN [1]
RP SEQUENCE OF 826-859 FROM N.A.
RC STRAIN=B6;
RX MEDLINE=21011895; PubMed=11129643;
RA Goltz M., Widen F., Banks M., Belak S., Ehlers B.;
RT "Characterization of the DNA polymerase loci of porcine
RT cytomegaloviruses from diverse geographic origins.";
RL Virus Genes 21:249-255(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=B6;
RA Widen F., Goltz M., Wittenbrink N., Banks M., Belak S., Ehlers B.;
RT "Characterization of the glycoprotein B gene of porcine
RT cytomegalovirus.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF268039; AAL47542.1; -.
DR InterPro; IPR00234; Glycoprotein B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprotein_B; 1.
SQ SEQUENCE 859 AA; 97797 MW; 8366A630ACEFF82F CRC64;

Query Match      23.9%; Score 1087.5; DB 12; Length 859;
Best Local Similarity 29.7%; Pred. No. 5.3e-60;
Matches 264; Conservative 193; Mismatches 325; Indels 107; Gaps 23;

Qy 11 VLALWLYQVALYS-----LSIAETGVTS-PNTATWSTESP-----LFGHYGTHDSS 57
Db 12 VLLINCYLVSSNSTTTSTTSASESVITGSPGTVSSETATPEVSTAFVNIITGNF----- 65
Qy 58 HGERGNENRDSSEONKNIYGSPTFPYRVCSASGVGDVFRPQTDHVCDD-ASDMVHSEG 116
Db 66 -----SOEYTEASDDEK-----YPRVCNMVAAGTDLRYFDNYITCNKNTYTQVSEG 112
Qy 117 ILLTYKNIIPFMRVKRYKRVVTTTSYVNGIYSDSITNOHTFYKSTPEWTEKMDTIYQ 176
Db 113 ILLLFKXNIIPHTFVRYTYTKELSFQTYRDVHVILVDRSSYKVPVDPVDEAGVNLNGQ 172
Qy 177 CFNSRLR-NTGGNLLTYVDRDDINMTVFLQPVGDVTPDKRYGSOPELYLEPGFWGYSYR 235
Db 173 CFSAAEIRNQGINRVYHKDDNTNTMRLYLKFGSINTNRYISTPEFQFTYGTWH-LYK 231
Qy 236 RRTVNCLEMDMFARSNPPDFPVTATGDTVEMSPWMSGDDDEHKNKHEKPMFVSVINNY 295
Db 232 SSSSINCIVTDTLAKSDYPYENFNLGTSVEISPFENGTS--KEVNEQMYEFSMKNNY 289
Qy 296 KV---VDYQN--RGTVPLGKTRIFLDREBYTLSEKHLKMSYCPFLWKAFYNGIOTEH 350
Db 290 TMLEKLDPEGNPKXTIP---TIAFLQKGDTLFSEVKEQNTNSHCKYTAWTAKKHALRADM 346
Qy 351 SGSYHVFANDITASFTTSKEDMKEFNTTYHCLNEEIKAEIEKKYA-KVNSTHSKYGDLKY 409
Db 347 TNSYHFMKMDTATMTVTKNTINLTGTGEYECVKNDIEKYITDTPQNKYNNTHNKNTENYSV 406
```

```

QY 410 FKTDGGLYLWQPLIQNRLLDAKNLNNETYSRRSRQAEESTTDPMMEMTNGAGGEYSS 469
Db 407 YETTTGLILWQPLVIRKSIKELKEFINETOHTYKR-----EIDG 446
QY 470 ENSITVAQVOYAYDNLRIRNNILEDLSKAWCREQRAALVMNELSKINPTSVMSIYNR 529
Db 447 SDSLVYASLQWYDALREYINAGFAQIAEAWCEDQKETNEVLSELAKISPSNVMSVIYDK 506
QY 530 PVSAKRIGDVISVNCIYVDQTSVLSLHSLRLI---SASDEKCFSPRPVTFKFWNDSTI- 585
Db 507 SLSAKLVGDAISVSSCVNWNQSTVKVHKDMRIYANGTANRETCSFRPVVTFEFSNNNSVQ 566
QY 586 -YKGLGQWNEILLTLYLTCENTEYFQAKTDMYIKNYEHLKTVPLSSITLDTFI 644
Db 567 QFTGLQPRNEILLGTIRVKECERNVKKVFFAGKEAYFFNYIYTKTVNISDINVVDTFI 626
QY 645 ALNFTLLENVDFKVIELYLDRDEKRLSNVFDIETMFREYNYAQRVSLGRKDLDDLSTNRN 704
Db 627 HLNKPLNTDFFVLRYMYSKNELAQAANFDLESILRDISYRSALYNIES---RIAPRKP 683
QY 705 QFVDAFGSLMDDGAGVQTVNAVSGVAT-LFSSIVTGFNFINKPFGG---MLMIIVVI 760
Db 684 DYVSGVDSFLHALGIGAPGGLGAALGATGAVTDFTLGIFSPFKNPFGLFSLMFLVLLV 743
QY 761 GVLFAIYFLTKTKIYETAPIKMIY-----EIDKLERECK-----797
Db 744 FLIFSVIY--KQNIY--TNVPGALFPYANSSGTVISNTHSYETNNKQEFENDRPDTS 800
QY 798 ---SEIAPISEEELERIVLAMHHQONSHMETKTRDKPKDSILTRAQNM 843
Db 801 NAVSEGSANKYSQEDAVCMMLMAIKLGDAYRRKNATKPSVLDKIRHL 849

RESULT 36
Q96910
ID Q96910 PRELIMINARY; PRT; 822 AA.
AC Q96910;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Glycoprotein B.
GN U39.
OS Human herpesvirus 7.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae.
OX NCBI_TaxID=10372;
RN [1]
RP SEQUENCE FROM N.A.
RA Hata A., Mukai T., Isegawa Y., Yamanishi K.;
RT "Identification and analyses of glycoprotein B of human herpesvirus
7.,"
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; D86577; BA13124.1; -
DR InterPro; IPR000234; Glycoprot B.
DR Pfam; PF00606; Glycoprotein_B_1.
DR ProDom; PD000693; Glycoprot_B_1.
SQ SEQUENCE 822 AA; 93178 MW; E3BFAA858362A619 CRC64;

Query Match 23.9%; Score 1086.5; DB 12; Length 822;
Best Local Similarity 33.6%; Pred. No. 5.8e-60;
Matches 272; Conservative 146; Mismatches 315; Indels 77; Gaps 24;

QY 78 GSSESTFFRYVCSAGVGDFVRFQTDHVCDD-ASDMVHSEGILYKNIIPFVRVKYR 136
Db 28 GHNOHLFPFICSIATGTDLVRFDFREVSCASVGSNIKTTEGILYIKTKIEAHTFSVRTPK 87
QY 137 KVTTSVYNGI---YSDSITNQHTFYKSIERWEIEKMDTIYQCNSLRLNTG--GNLL 190
Db 88 KEITFQTYRDVGTVFLDRTVT--TLPMPIE--EVHMVNTAEARCUSSISVKSSEEEYV 143
QY 191 TYVDRDDINMTVFLQPVGDVTPDVKRYGSOPELYLEPGFWGYSRRRTTVCNCLMDMFAR 250

```

```

Db 144 AVHKDEYVNVKTLDLIPLNFKSDTVRRYITTTKEPFLRNGPLW-FYSTSTSIINCIVTDCIAK 202
QY 251 SNPPPPFFVTATGDTVEMSPFWSGEDDHE-NKMHKXPWFVSVINNY---KVVDYQNRGTV 306
Db 203 TKYPDFPFALSTGETVEGSPFPYNGINSKTFNEPTEKILF---RNNYTMKLTFPDGGSKGNF 259
QY 307 PLGKTRIFLDREBYTLSEWEKHLKNMYSYCLTLTWKAFYNGIQTSHSGSYHYFVANDITASFT 366
Db 260 VILTKWAFLEKNGNTIFSNEVQNEESSICLLKHWMTIPHALRAENANSFHEIAELTASFV 319
QY 367 TSKEDMKFNTTYCHLNEEBIKABIEKKY-AKVNSTHSKYGDLYKFKTDGGLYLWQPLIQ 425
Db 320 TGSNVTLSDSKYNCINSVNTSILDEIYQTYQVNSHDKNGSYEIFKTEGDLILWQPLIQ 379
QY 426 NRLDDAKNLNNETYSRRSRQAEESTTDPMMEMTNGAGGEYSSENSITVAQVOYAYDNL 485
Db 380 RKLTVLEN-FSNASRKRKR-----ELETNKDIYVVLQYLYDYL 418
QY 486 RIRINNILEDLSKAWCREQRAALVMNELSKINPTSVMSIYNRPVSAKRIGDVISVNC 545
Db 419 KOYINTALGLAEAWCLDQKRTITVLHLSKISPSGIIISAVYKPKMSAKLIGDVLAVSKC 478
QY 546 IYVDQTSVLSLHSLRLLSASDE---KCFSPRPVTFKFMNDS-TIYKGLQGVNNEILLTTT 601
Db 479 IEVNTQTSVOLHKSMLRTKDSYDTLRCYSRPLLTYSFANSKETYLGQLGLDNEILLGNH 538
QY 602 YLETC-QENTEYFQAKTDMYIKNYEHLKTVPLSSITLDTFIALNFTLLENVDFKVI 660
Db 539 RTEECQSNKTFLSGKF-AHIFKDYTYVNSSLITEIALDAFVDINIDPLENADFTLLE 597
QY 661 LYTRDEKRLSNVFDIETMFREYNYAQRVSLGRKDLDDLSTNRNQFVDAFGSLMDDLGAV 720
Db 598 LYTDELSKANVDFLETLIREVNSYK---SALHETKATVTVTPTVIGGIDTFFKGLGAL 654
QY 721 GQTVNAVSGV-ATLFSSIVTGFNFINKPFGGMLMIIVVIGLFAIYFLTKTKIYETA 779
Db 655 G-LGLGAVLGTAGALGDVGVNGVSPFNKPNPFGGALTILLTGLVIGLVIPLFLRHKRLAQT 713
QY 780 PKMIYPEIDKUKER--EGKSEIAPISEEELERIVLAMHHQONSHMETKTRD--PKDS 835
Db 714 PIDILFPYTSKTSNSTVQLQTSVQAQVKEPLD-----SSPYLTKTKDTEPQGD 762
QY 836 ILTRAQNMRLKRSYSNLKNABSVEMLNTL 865
Db 763 DITHNE-----YSQV---EALXMKLAI 782

RESULT 37
O56280
ID O56280 PRELIMINARY; PRT; 822 AA.
AC O56280;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Envelope glycoprotein gB.
GN U39.
OS Human herpesvirus 7.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae.
OX NCBI_TaxID=10372;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RR;
RX MEDLINE=98240941; PubMed=9581785;
RA Megaw A.G., Rapaport D., Avidor B., Frenkel N., Davison A.J.;
RT "The DNA sequence of the RK strain of human herpesvirus 7.,"
RL Virology 244:119-132(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=RR;
RA Megaw A.G., Frenkel N., Davison A.J.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF037218; AAC40753.1; -

```


Glycoprotein B (Fragment).

GB.
Caprine herpesvirus 2.
Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
Gammaherpesvirinae.

NCBI_TaxID=135102;

(1)

SEQUENCE FROM N.A.
MEDLINE=21212029; PubMed=11311431;
Chmielewicz B., Goltz M., Ehlers B.;
"Detection and multigenic characterization of a novel gammaherpesvirus
in goats.";
Virus Res. 75:87-94(2001).
EMBL; AF283477; AAK26669.1; --
InterPro; IPR00234; Glycoprot.B.
Pfam; PF00606; Glycoproteins_B; 1.
ProDom; PD000693; Glycoprot_B; 1.
NON TER 1
SEQUENCE 349 AA; 39749 MW; F8C18988E1B42B8B CRC64;

Query Match 23.9%; Score 1084.5; DB 12; Length 349;
Best Local Similarity 59.5%; Pred.No.2.3e-60;
Matches 207; Conservative 61; Mismatches 65; Indels 15; Gaps 3

Qy 524 SMYINRPVS AKRIGDVISVSNICVDQTSVLHKSRLLSASD-EKCFSRPPVTFKMND 582
Db 1 SSIYGRPVSAKRIGDVISVSHCVVDDQQSVSLHRNMRVPVKDHAHECYSPPTPKFIND 60
Qy 583 STYIKQLGVNNBILLTTTLETCQNTETYYFOAKTDWYIKNYEHLKTVPLSSITLDT 642
Db 61 SHLYKQLGVNNBILLTTTALEYCHENTHEYFGGDNMVFYKNRYHVKTMPVTVALDT 120
Qy 643 FIALNFTLLNVDFKVIELYTRDEKLSNVFDIETMFREYNYAQRVSGLRKDLDDLSTN 702
Db 121 FIVLNUTLVENIDFQVIELYSREKKRMSTVFDIETMFREYNYTORVTLGLRRDLTDIKSN 180
Qy 703 RNQVDAFGLMDLDLGAVGTQVVNAVSGVATLFSSIVTGFINFIKNPFPGMLMIIVIGV 762
Db 181 RNQVDAFGLMDLDLGAVGTQVVNAVSGVATLFSSIVTGFINFIKNPFPGMLLFLGLAAV 240
Qy 763 LFAIFYELTKTKIYETAPIKMIIPEIDKLKEREGSEIAPISEEBLERIVLAMHIHQNS 822
Db 241 VVMVILMRRAASFQAQNPQMIVPDITQITKOREVNMMVEPISKQELDIRMLAMHDYHNK 300
Qy 823 HMETKTKRKPKDS-----ILTRAONMLRKSGSYNLNAESVE 860
Db 301 ----QEKEDQBSGGGFLNRGDWLDKAKNLRKAGYOPLKRTDSSE 344

RESULT 39 -
Q8V429 PRELIMINARY; PRT; 851 AA.
AC Q8V429;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Glycoprotein B (Fragment).
OS Porcine cytomegalovirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=109993;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=P1;
RX MEDLINE=21636973; PubMed=11778702;
RA Widen F., Goltz M., Wittenbrink N., Ehlers B., Belak S.;
RT "Identification and sequence analysis of the glycoprotein B gene of
porcine cytomegalovirus";
RL Virus Genes 23:339-346(2001).
DR EMBL; AF394057; AAL40232.1; --
DR InterPro; IPR000234; Glycoprot.B.
DR Pfam; PF00606; Glycoproteins_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.

DE InterPro; IPR00234; Glycoprot_B.
DR Pfam; PF00606; Glycoproteins_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
SQ SEQUENCE 822 AA; 93148 MW; AF79DB32F9DE6715 CRC64;

Query Match 23.9%; Score 1086.5; DB 12; Length 822;
Best Local Similarity 33.6%; Pred.No.5.8e-60;
Matches 272; Conservative 146; Mismatches 315; Indels 77; Gaps 24;

Qy 78 GSPSTFYRVCSAGVGDFRFQTHVCPD-ASDMVHSGILLIYKONIIPMERVRKYR 136
Db 28 GHNQHPFRICSTATGDLVRFPREVSCASYGSNIKITTEGILLIYKIEAHFVSURIFK 87
Qy 137 KVVTISVYNGI----YSOSITNHQTFFYKSIEPWETERKDITTYQCNSLRNLNTG--GNLL 190
Db 88 KELTFQTTVDVGVYFLDRTVI--TLWPPIE--EVHMVNTPEARCLSSISVRSSEEEV 143
Qy 191 TYVDRDINNVTFLQPVDGTPVKYGYGOPELYEPGFWSYRRRTVNCNELMDMAR 250
Db 144 AYKHDEYVNKTLDLIPLNFKSDTVRRYITKEPFLRNGPLW-FYSTSINSIVDCIAK 202
Qy 251 SNPPFPFFVATGDTVMESPFWSGEDDHE-NKMHKEKPFVSVINNY--KVVDYONRGTV 306
Db 203 TKYPFFEFALSTGETVEGSPFYNGINSKFNEPEKILF---RNNVTMLKTDDGSGNF 259
Qy 307 PLGKTRIFLDREYTLISWEKHLKNMSYCPLTLWKAFYNGIQTESHSYHFVANDITASFT 366
Db 260 VTLTWKMAFEKGNTITSWEQNEESSICLLKHWMTIPHALARAEANANSPHFIAQELTASPV 319
Qy 367 TSEDMEKFNWTTTHCLNEERIKABIEKY-AKNSTHSKYGDLYKYPTDGGLVLMQPLIQ 425
Db 320 TGKSNYTLSDSKYNCSINYSYTLDEIYOTVNNSHDKNGSYEIFKTEGDLILIWOPLIQ 379
Qy 426 NRLLDAKNLNNETYRRSRROAESTTDPMMEMTNGAGGESSENSITVAOVAYDNL 485
Db 380 RKLTUVLEN-FSNASRRKR-----ELETKNDIVYVOLXYDYL 418
Qy 486 RIRINNIEDLSKAWCREQHRAALVWNELSKINPTSVMSIYNRPVSAKRIGDVISVNC 545
Db 419 KDYINTALGKLAECWLCDKRTITVLHELKISPSGIISAVYGPMSAKLIGDVLAISK 478
Qy 546 IVDQTSVSLHKSRLLSASDE---KCFRRPPVTFKMND-S-TYKQLGVNNELLTTT 601
Db 479 IEYNQTSVQLHKSMRLTKDSSYDALRCYSRPLITYSFANSSKETVYLGQLGLDNETLLGNH 538
Qy 602 YLET-C-QENTEYVFOAKTDWYIKNYEHLKTVPLSSIITLDTFIALNFTLLNVDFKVI 660
Db 539 RTECQSNTKIPLSGKF-AHIFKDYTYVNSSLIETEIALDAFDVLNIDPLENADFTLLE 597
Qy 661 LYTRDEKRLSNVFDIETMFREYNYAQRVSGLRKOLLIDLSTNRNOVFDAFGLMDLDLGAV 720
Db 598 LYTKDELKANVDFLETLIREYNSYK--SALHHIETKIATVTPYIGDITFFKGLGAL 654
Qy 721 GQTVNAVSGV-ATLFSIVTGFINFIKNPFPGMLMIIVIGVLFALYFLTKTKIYETA 779
Db 655 G-LGLGAVLGTATAGLDGVVGVFSPKPNPFGALTILTLLGLVIFLFLRHRLAQ 713
Qy 780 PIKMIYPEIDKLKER--ECKSETAPISEEBLERIVLAMHIHQNSHMETKTKD--PKDS 835
Db 714 PIDLPFYTSKTSNSVLQATQSOAOVKPELD-----SSPPYKTNKTKOTEPOGD 762
Qy 836 ILTRAQNMLRKRSYGSYNLNAESVEMNLTL 865
Db 763 DITHTNE-----YSQV---EALKMKAI 782

RESULT 38
Q8VQ00 PRELIMINARY; PRT; 349 AA.
AC Q8VQ00;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)

```
FT NON TER 1 1
FT NON TER 851 851
SQ SEQUENCE 851 AA; 96956 MW; C7E2404646F5A66D CRC64;

Query Match 23.8%; Score 1080.5; DB 12; Length 851;
Best Local Similarity 29.6%; Pred. No. 1.4e-59;
Matches 263; Conservative 193; Mismatches 326; Indels 107; Gaps 23;

Qy 11 VLALWYLQVALYS-----LSIAETGVTS-PNNTATWSTESP-----LTGHYGHDSHSS 57
Dy 11 VLLINYGLVSSNSTTSTTSASESVTTGSGTTSATPSTAFVNITGNF----- 64
Qy 58 HGERGNENRDEBQNKNIYSPSTFPYRVCSAGVGDFRQFDHVCPO-ASDMVHSEG 116
Dy 65 -----SOEYTPASDEK-----YFPRVCNMAVGTDLYRFDNYITCNKYNITETQISEG 111
Qy 117 ILLIYKONIIPFMRVRKRYKVVTTSTVYNGIYSDSITNQHTFYKSPETEMEKMDIYQ 176
Dy 112 ILLLFKKNIVPHTFFVRYTKELSFQTYRDVHVILVDRSSYKVPVPVDEAGVINLNGQ 171
Qy 177 CFNSLR-L-NTGNNLLTYVDRDDINMTVLPQVDGVTDPVKRYGSOPELYLEPWFNGSYR 235
Dy 172 CFSAAEIRNQGINRVYHKDDNTNMTMLRYLKFSGSTINTRYISTPBFQTYGTHW-LYK 230
Qy 236 RRTVNCCLMDMFARSNPPDFPVATGDTVEMSPFWSGEDDHENKMKHPFVSVINNY 295
Dy 231 SSSSINCIVDTLAKSDYPYENFLGTGESVEISPPFNGTS--KEVNEQMYFPMKNY 288
Qy 296 KV--VDYQN--RGTVPGLGTRIFLDREEVYTLSEKHLKMSYCPLTLWKAFYNGIOTEH 350
Dy 289 TMLEKLEDPNGPKTIP--TIAFLQKGDTLFSWEVKEQTNCHCKYTAWTAKHIALRADM 345
Qy 351 SGSHVFVANDITAGFTSKEDMKFNTYHCLNEEIKAEIEKKYA-KVNSTHAKYGDLYK 409
Dy 346 TNSYHFMMKDDTATMTVTTNTINLTGTEYECVKNDIEKITDTPQNKYNNHTKNENYSV 405
Qy 410 FKTDGGLYLVQPLQNRLLDANKLNKNTYVSRSRQAESTTDPMMEMTNGAGGEYSS 469
Dy 406 YEITTGILLVQPLVIRKSIKELKEFINETQHYKR-----EIDG 445
Qy 470 ENSITVAQOVAYDNLIRINNILEDLSKAWCREQHRRAALVWNLKSKINPTSVMSIYNR 529
Dy 446 SDSLVYASLOMYDALREYNAGPAQAEAWCEDQKRTNEVSELAKISPSNVMSVIYDK 505
Qy 530 PVSAKRIGDIVISVNCIVDQTSVLSHLSRL--SASDEKCFSPRPVTFKPMWDSIT- 585
Dy 506 SLSAKLVGDALISVSSCVNVNQSTVKVHKDMRIYANGTANRETCTFSPRVVTFEFSNNNSVQ 565
Qy 586 -YKQGLGVNNEILLTLYLETQENTEYVFOAKTDMYIKNYEHLKTVPLSSITLDTFI 644
Dy 566 QFTQGLGPRNEILLGTVREKERNVSVKVPFAGKEAYFFNYIYTKTVNISDINVVDTFI 625
Qy 645 ALNFTLLENVDFKVIELYTRDEKRLSVDFDIETMFREYNYAQRVSGRLKDLDLSTNRN 704
Dy 626 HLNIKPLENTDFEVLRYMYSKNELAQANFDLESRLDRINSYRSALYNIES--RIAPRKP 682
Qy 705 QFVDAFGLSMDLDGACQGVVNAVSGVAT-LFSSIVTGTFNFKNPGG---MLMIIVVI 760
Dy 683 DYVSGVDVSFLHALGIGAPGLGALGMATGAVTDFTLGTGIFSFKNPFGGLFMSLFFVLV 742
Qy 761 GVLPAIYFLTKTKIYETAPIKMIYP-----EIDKLKEREK----- 797
Dy 743 FLIFSVY--RQKNIY-TNPVGAFLPYANSSGTVINSHSYETNNKQEFENDRKPDT 799
Qy 798 ---SETAPISEEBELERIVLAWHIHQONSHMETKTRKDPKDSILTRAQNM 843
Dy 800 NAVSEGSANKYSQEDAVCMLEMAIKNLGDAYRRKNATKPSVSLDKIRHL 848

RESULT 40
Q9DKT9
ID Q9DKT9 PRELIMINARY; PRT; 845 AA.
AC Q9DKT9;
```

```
QY 568 -----KCFSRPVPVTFKFMND--STIYKQGVNNBILTTTLYETCQENTYFYFOAKTDM 620
Db 546 WGRDMQCYSRPLVTRFLDNETASTIRTGOLGVNDNEILLGNRYTELCCQENSIRYFVAGAQI 605
QY 621 YIYKNYHLKTVPLSSITLTDTFFIALNTLLENVDFKVIELYTRREKLSNVDFDIETWFR 680
Db 606 HVFQDYDFYHTIKLSDDVIDTFFVHLNLSFLQNDIFQMLRLYTOBEQYASRLLDLETLRL 665
QY 681 EYNYAQRVSGLRKDLLDLSTNRNDFVDAFG--SLMDDLGAQGVQVNVAVSGVATLFPSSI 738
Db 666 DENTYRQRIYKLEQAIV-----TKPYVPAGMQQALQSLGSGV-----SVITGLCAMQSL 716
QY 739 VTGFNFINKPFGGMLIIWIGVLFAIYFLTKTKIYETAPIKMIYPEIDK-LKEREGK 797
Db 717 VSGVASFLQNPFGGTLIIIGCIIVGIIYNNRNQSRGSPIDYFVFPVQNLTPQRLQ 776
QY 798 SEIA--PISEEL-----ERVLAMHIIHQONSHMETK-----TRKDPKDSILTRAQN 842
Db 777 QHVGPPSYDESIGSHTYSKEDALLMLKAMKELDKSEKAQIEATKSQP--SIIDRI-- 832
QY 843 MLRKRSYGNLKN 855
Db 833 ---RRRGYITLSS 842

RESULT 41
QY Q8V430 PRELIMINARY; PRT; 860 AA.
AC Q8V430;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Glycoprotein B.
OS Porcine cytomegalovirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=109993;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1469;
RX MEDLINE=21636973; PubMed=11778702;
RA Widen F., Goltz M., Wittenbrink N., Ehlers B., Banks M., Belak S.;
RT "Identification and sequence analysis of the glycoprotein B gene of
RL porcine cytomegalovirus.";
RL Virus Genes 23:339-346(2001).
DR EMBL; AF394056; AAL40231.1; -.
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein B; 1.
DR ProDom; PD000693; Glycoprot B; 1.
SQ SEQUENCE 860 AA; 97911 MW; 36781369491D1677 CRC64;

Query Match 23.4%; Score 1063; DB 12; Length 860;
Best Local Similarity 29.2%; Pred. No. 1.9e-58;
Matches 260; Conservative 193; Mismatches 330; Indels 106; Gaps 23;

QY 11 VLALWLYQVALYSLISIA---ETGVTS-PNPTATWSTESP-----LTGHYGTDDSS 57
Db 12 VLLINYGLVSNSTTTATAPESVTTGSSGTSSTETAPSTAVFNITGNF----- 65
QY 58 HGERGNENRDESEQKNKIYSPSTFFPYRVCSAGSGVDVFRQTDHVCVD-ASDWHVSEG 116
Db 66 -----SQEYTEASDDEK-----YPERVCNMAVGTDLYRFDNYITCKYDTEQYSEG 112
QY 117 ILLYKQNIIPFMFRVRYKRVKVTSTVYNGIYSDSITNQHTFYKSIPEWTEKMDTIYQ 176
Db 113 ILLFPKNIVPHTFFVRYTKLSQTTYRDVHVILYVDRSYYKVPVPVDEAGVINSNGQ 172
QY 177 CFNSLRL-NTCGNLLTYVDRDINNTVFLQPDVGPDPVKYSGOPELYLEPGFWGSYR 235
Db 173 CFSAAEIRNOGINRVYVHKDDNTNNTMLRYLKLFGSTINTRYSTPEQFYGTGHW-LYK 231
QY 236 RRTTVNCELMDMFARSNPPDFDFVTATGDTVMSPFWSGEDDHENKMKHKEKWPVSVINNY 295
```

```
Db 232 SSSSINCITVDTMGKSDYPYDNFILGTGESVEISPFNGTS--KEVNEQMYFYSMKRNY 289
QY 296 KV---VDYQN--RCTVPLGKTRIELDREYYTSLWEKHLKNMSYCLPLTLWKAFYNGIQTEH 350
Db 290 TMLSEKLDPEGNPKKTIIP---TIAFLQKGDTLFSWEVKEQTYNYCKYTAWTKKHHLRADM 346
QY 351 SGSYHFVANDITASFTTSKEDMKEFNTTYHCLNEEIKAEIEKKY-AKVNSTHSKYGDILKY 409
Db 347 TNSYHFMMKOWTATMTTKNTINLTGTIEYECVKNDIEKYITDTFQSKYNNTHNKTENYSV 406
QY 410 FKTGGGLVWQPLIQNRLLDAKNLANETVSRRSRROAEBSTTDPMMEMTNGAGGESS 469
Db 407 YETTTGLILWQPTVRKSIKELKEFINETOQHTYKR-----ETDG 447
QY 470 ENSTVAVQVQYAYDNLAIRINNILEDLKAWCREQHRRAALVWNELSKINPTSVMSMIYNR 529
Db 448 SESLVYASLQWYDALREYNAGFAQAEAWCEDQKRTNEVLSELAKNPSNVMSVYDK 507
QY 530 PVSARIGDVISNCCIWVQTSVSLHKSRLLL---SASDEKCFSRPPVTFKFMNDSTI- 585
Db 508 SLSAKLVGDALSVCVNVNQSTVKVHKDMRIYANGTANRETCTCFRPPVTFEFSNNSVQ 567
QY 586 -YKQQLGVNNEILLTTTYLETQCENTEYFYFOAKTDMYIYKNYHLKTVPLSSITLTDTFI 644
Db 568 QFTQGLGRNEILLGTHRVEKCCERNVSVKVPFAGKEAYFFYNYIYTKTVNISDINVDTFI 627
QY 645 ALNFTLLENVDFKVIELYTRREKLSNVDFIETMFRYNYAQRVSGLRKDLLDLSTNRN 704
Db 628 HLNKPLENTDFEVLRYMYSKNELAQANIFLESLLIDINSYRSALYNIES---RIAPRKP 684
QY 705 QFVDAFGSLMDDLGAQGVQVNVAVSGVAT-LFSSIVTGFNFINKNPFPG---MLMIIVVI 760
Db 685 DYSGVDSFLHALGIGAPGGLGAALGNATGNTDFTLGTGIFSFKNPFGGLFSLFFVLV 744
QY 761 GVLFAIYFLTKTKIYETAPIKMIYTP-----BIDLKEREKG----- 797
Db 745 FLIFSYY--RQKNYI-TNPVGALFPVANSSGTVISNTHSYETNNKQSENDERKPDTS 801
QY 798 ---SEIAPISEEELERIVLHAMHIHQONSHMETKTRKDPKDSILTRAQNM 843
Db 802 NAVSEGSANKYSQEDAVCMMLMAIKNLGDAYRRKNTTRKPSVLDKIRHL 850

RESULT 42
QY Q8V614 PRELIMINARY; PRT; 859 AA.
AC Q8V614;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Glycoprotein B.
GN GB.
OS Porcine cytomegalovirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=109993;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=55b;
RX MEDLINE=21011895; PubMed=11129643;
RA Goltz M., Widen F., Banks M., Belak S., Ehlers B.;
RT "Characterization of the DNA polymerase loci of porcine
RL cytomegaloviruses from diverse geographic origins.";
RL Virus Genes 21:249-255(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=55b;
RA Widen F., Goltz M., Wittenbrink N., Banks M., Belak S., Ehlers B.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF268040; AAL47541.1; -.
DR InterPro; IPR000234; Glycoprot B.
DR Pfam; PF00606; Glycoprotein B; 1.
```

```

DR ProDom; PD000693; Glycoprot B; 1.
SQ SEQUENCE 859 AA; 97912 MW; 6B9C023F79CB5B2B CRC64;

Query Match      23.3%; Score 1061; DB 12; Length 859;
Best Local Similarity 30.9%; Pred. No. 2.5e-58;
Matches 248; Conservative 177; Mismatches 304; Indels 74; Gaps 19;

Qy 11 VLALWLYQVALYSLSLA-----ETGVTSPPN-----TATWSTESPLTGHVTHDSSHGE 60
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 71
Qy 61 RGNNEIRDSEBQNKNIYGSPTPPYRVCSAGVDVFRFQTDHVCPPD-ASDMVHSEGILL 119
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 115
Qy 120 IYKNIIPFMRVKYKVVTTSTVNGIYSDSIHQHTFYKSIETPEKMDTIYQCFN 179
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 175
Qy 116 LFKNIIVPHFTFFVRYTKELSFQTYRDVHVILVDRSSYKVPVVDVDEAGYINLNGCFS 238
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 234
Qy 180 SLRL-NTGNNLLTYVDRDDINMTVFLQPDVGVTPDKRVGSQPELYLEPGWFGVSRRT 297
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 292
Qy 239 TVNCELMDMFARSNPPDFVATGDTVEMSPWGSDEDDHENKMKHPWFSVNNYKV- 353
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 349
Qy 235 SINCIITDTMGKSDYPYDNFILGTGSVEISPFENGTS--KEVNEQMYFYSKRNVTML 412
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 409
Qy 298 --VDYQN--RGTVPLGKTRIFLDREBYTLSWEKHLKMSYCPLTLKAFYNGIQTEHSGS 472
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 450
Qy 293 EKLDPENGPKTIP---TIAFLQKGDTLFSEWVKQYNTNRYLKFEGSTINTRIYSTPDTFQTYGTHW-LYKSSS 532
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 510
Qy 354 YHFVANDITASFTTSKEDMKFNTTYHCLNEEIKAEIEKKY-AKVNSTHSGYDGLKYFT 587
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 570
Qy 413 DGGYLWQPLIQNRLLDANKLNNETYSRRSRQAESTTDDPMWMTGNGAGGEYSSENS 647
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 630
Qy 473 ITVAQVOYAYDNLRIIRNNILEDSLKAWCREQRAALVWNLKINPTSVMSIYNRPVS 707
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 687
Qy 511 AKLVGDAISVSSCVNQSVKHKDMRIYANGTANRETCSRPVTFEFSNNVQOFT 763
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 747
Qy 588 GOLGVNNEILLTTLTYLETQENTYEFYQAKTDMYIYKNYEHKLTVPVLSITLDTFI 786
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 767
Qy 648 FTLEENVDFKVIELYTRDEKRLSNVFDIETMFREYNYAQRVSGRLKDLDDLSTNRNQP 860
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 860
Qy 631 IKPLENTDFEVLRMYSKNELAQANIFDLESLLDINSYRSALYNIES---RIAPRPDYV 860
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 860
Qy 708 DAFGLMDDLCAVQGVVNAVSGVAT-LFSSIVTGFNFINKNPGG---MLMIIVIGVL 860
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 860
Qy 688 SGVDSFHALGIGAPGGLGAALGNATGAVTDLTGTFESFPKPNPGLFSLMFLVLLVLI 860
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 860
Qy 764 FAIYFLTKTKIYETAPIKMIYP 786
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 786
Qy 748 FSVYY--RQKNIIY-TNPVGALFP 767
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 767

RESULT 43
Q8V613 PRELIMINARY; PRT; 860 AA.
AC Q8V613;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Glycoprotein B.
OS Porcine cytomegalovirus.

```

```

OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=109993;
RN [1] SEQUENCE FROM N.A.
RP STRAIN=OF-1;
RC MEDLINE=21011895; PubMed=11129643;
RA Goltz M., Widen F., Banks M., Belak S., Ehlers B.;
RT "Characterization of the DNA polymerase loci of porcine
RT cytomegaloviruses from diverse geographic origins.";
RL Virus Genes 21:249-255 (2000).
RN [2] SEQUENCE FROM N.A.
RP STRAIN=OF-1;
RC Goltz M., Widen F., Banks M., Belak S., Ehlers B.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF268041; AAL47540.1; -.
DR InterPro; IPR000234; Glycoprot_B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot_B; 1.
SQ SEQUENCE 860 AA; 97834 MW; E61F989F175BF433 CRC64;

Query Match      23.3%; Score 1059; DB 12; Length 860;
Best Local Similarity 29.1%; Pred. No. 3.3e-58;
Matches 259; Conservative 194; Mismatches 330; Indels 106; Gaps 23;

Qy 11 VLALWLYQVALYSLSLA---ETGVTS--PPNTATWSTESP-----LTGHVTHDSS 57
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 65
Qy 58 HGERGNNEIRDSEBQNKNIYGSPTPPYRVCSAGVDVFRFQTDHVCPPD-ASDMVHSEG 116
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 112
Qy 117 ILLIYKONIIPMFRVKYKVVTTSTVNGIYSDSIHQHTFYKSIETPEKMDTIYQ 176
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 172
Qy 177 CENSLRL-NTGNNLLTYVDRDDINMTVFLQPDVGVTPDKRVGSQPELYLEPGWFGVSYR 235
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 231
Qy 236 RRTTVNCELMDMFARSNPPDFVATGDTVEMSPWGSDEDDHENKMKHPWFSVNNY 295
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 289
Qy 296 KV---VDYQN--RGTVPLGKTRIFLDREBYTLSWEKHLKMSYCPLTLKAFYNGIQTEH 350
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 346
Qy 351 SGSYHFWANDITASFTTSKEDMKFNTTYHCLNEEIKAEIEKKY-AKVNSTHSGYDGLKY 409
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 406
Qy 347 TNSYHFMKMDTAMTWTTKNTINLTGTEYECVKNDIEKYITDTFQSKYNNTHNKTENYSV 469
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 447
Qy 410 FKTGGLYLWQPLIQNRLLDANKLNNETYSRRSRQAESTTDDPMWMTGNGAGGEYS 469
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 447
Qy 470 ENSITVAQVOYAYDNLRIIRNNILEDSLKAWCREQRAALVWNLKINPTSVMSIYNR 529
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 507
Qy 448 SESLVYASLQMYDALREYNAGFAQIAEAWCEDQKRTNEVLSELAKINPSNVMVIYDK 507
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 507
Qy 530 PYSAKBIGDIVSNCIVVDQTSVLSLHKSIRLL---SASDEKCFSPPTVTFEKNWDSIT- 585
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 567
Qy 508 SUSAKLVGDAISVSSCVNQSVKHKDMRIYANGTANRETCSRPVTFEFSNNVQ 567
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 567
Qy 586 -YKGLGVNNEILLTTLTYLETQENTYEFYQAKTDMYIYKNYEHKLTVPVLSITLDTFI 644
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 627
Qy 568 QFTGQLGPRNEILLGTHRVKCEKERNVKKVPFAGKEAYFFVNYIYTKTVNISDINVDTFI 627
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 627
Qy 645 ALNFTLLENVDFKVIELYTRDEKRLSNVFDIETMFREYNYAQRVSGRLKDLDDLSTNRN 704
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 684
Qy 628 HLNKPLENTDFEVLRMYSKNELAQANIFDLESLLDINSYRSALYNIES---RIAPRPK 684
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 684

```

```

QY 705 QFVDAFGSLMDLGAAGVGTGVNAVSGVAT-LFSSIVTGFINFIPKPPFGG---MLMIIVVI 760
Db 685 DVVSGVDSDHALGIGAFGGLGAALGMATGAVTDLTGIFSFKNPFGGLFSMLFFVLV 744
QY 761 GVLFYAIPLTKTKIYETAPIMYIP-----EIDKLKEREGK----- 797
Db 745 FLIFSVY--RQKNY--TNPVGPALFPYANSSSGTVISNTHSYETNNKQESENDRKPDTS 801
QY 798 ---SRIPISEBELERIVLANHIIHQONSHMETKTRKDPKDSILITPAQNM 843
Db 802 NAVSEGSANKYSQEDAVCMWMAIKNLGDAYRRKNTTKPSPVLDKIRHL 850

RESULT 44
Q9WT25 PRELIMINARY; PRT; 830 AA.
AC Q9WT25;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Glycoprotein B.
GN U39.
OS Human herpesvirus 6.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Roseolovirus.
OX NCBI_TaxID=10368;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HST;
RA Isegawa Y.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=HST;
RX MEDLINE=99412319; PubMed=10482554;
RA Isegawa Y., Mukai T., Nakano K., Kagawa M., Chen J., Mori Y.,
RA Sunagawa T., Kawanishi K., Sashihara J., Hata A., Zou P., Kosuge H.,
RA Yamaniishi K.;
RT "Comparison of the complete DNA sequences of human herpesvirus 6
RT variants A and B.";
RL J. Virol. 73:8053-8063(1999).
DR EMBL; AB021506; BAA78260.1; -.
DR InterPro; IPR00234; Glycoprot B.
DR Pfam; PF00606; Glycoprotein B; 1.
DR ProDom; PD000693; Glycoprot B; 1.
DR PROSITE; PS00044; HTH LYSR FAMILY; 1.
SQ SEQUENCE 830 AA; 93136 MW; 6DBCCF3PFC3157B CRC64;

Query Match 23.1%; Score 1052.5; DB 12; Length 830;
Best Local Similarity 32.4%; Pred. No. 8.1e-58;
Matches 245; Conservative 157; Mismatches 295; Indels 59; Gaps 15;

QY 83 FPRVCSAGVDVPRFOTDHCV-PDASDMVHSEGILLIYQNIIPFMRVRKYKVVTT 141
Db 36 YPFRICSTAKGTDLRFRDIDSCSPYKGNKMSGFFIYKTIETIYTPFVRTYKNELTF 95
QY 142 STVYNGIYSDSTINQHTFYKSIETPEKMDTIYOCNSLRN--TGCNLLTYVDRDDIN 199
Db 96 QTSYRDVGTVFLDRTVMGLAMPYEAANLVNSRAOCYSABAIVAKRPGDVTFAYHEDNNKN 155
QY 200 MTFVLQPDVGPDKRYKGSOPELYEPFGWGSYRRRTTVNCELMDFARSNPPFPFV 259
Db 156 ETLELFLPLNFKSVTKRITTKPEYFARGPLW-LYSTSTSLNCIVTEATAKAPFFSYFA 214
QY 260 TATGDTVMSFPWSECD-DHKNMHEKFWFVSINNYKV--DYQN--RGTVPLGKTRIFL 315
Db 215 LTTGIVSGSPFFDGSNGKHAEPLEK--LTIENYTMIEDLNMGMNGATTLVKRIAPL 271
QY 316 DREEVTLNWEKHLKNMVCPLTLKFAFYNGIQTEHSGSHFVANDITASFTTSKEDMKEF 375
Db 272 EKGDTLFSWEIKEENESVCMUKHTVTTHGURAEQDETHYFISKELTAFTPKDLSNLT 331
QY 376 NTTVHCLNEEIKAEIEKRY-AKVNSTHSGYGLKFKYKTDGGGLYLWQPLIQNRLLDANK 434

```

```

Db 332 DPKQCTIKNEFENIIEKVMYSNDYNTYSMNGSYQIFKTTGDLILIMQPLVOKSLM---V 387
QY 435 LANNETYSRRRQRAESTTDPMMENTGAGGEVSSSENSITVAQVOYAYDNLIRINNILE 494
Db 388 LEQGSVNLRRRLDV-----DVKSRHDILYVQOQYLYDTLUDKOVINDALG 431
QY 495 DLSKAWCREQHRALVNNELSKINPTSVMSIYNRPVSAKRIGDVISVSNCIIVVDQTSVS 554
Db 432 NLAESWCLDQKRTITMLHELKSPSSIVSEVYGRPISAQLHGDVLAISKIEVNSQSVQ 491
QY 555 LHKSRLRLSA----SDEKCFSPVPTFKWNDS--TIYKGQGVNNEILLTYYTLETQEN 609
Db 492 LHKSMRVVDKGLRSETVCYNRPVLPVTFSPVNSTPEVVPVQGLDNEILLGHRTECEIP 551
QY 610 TEYVFOAKTDMYIYKNYEHKLTVPVLSSTITLDTFIALNLTLENNVDKVELYTRDEKRL 669
Db 552 STKIFSGNHAHVYTDYTHNSTPIEDIEVLDAPFIRUKIDPLENADFKLDLJLSPDELSR 611
QY 670 SNVPDIETMFREYNVYQVRVGLRKLDDLSTNRNQFVDAFGSLMDLGAAGVGTGVNAVVS 729
Db 612 ANVEDLENILREYNSYK--SALVTIEAKIATNTPSYVNGINSFLQGLGAIQTGLGVSIS 668
QY 730 GVATLFSIVTGFINFIPKPPFGMLIIVIGVLFYAIYFLTKTKIYETAPIMYIP--- 786
Db 669 VTAGALGDIVGGVSVFLKNPFGGLMLILAIIVVVIIIVFVRQKHVLSKPIDMMFPYAT 728
QY 787 -----EIDKLKEREGKSEIAPISSE 807
Db 729 NPVTTSVSVTGTATAVKTPTSVKRDADGGTSA-VSEKE 763

RESULT 45
Q8QS33 PRELIMINARY; PRT; 921 AA.
ID Q8QS33
AC Q8QS33;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 23, Last annotation update)
DE Envelope glycoprotein UL55.
OS Chimpanzee cytomegalovirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=188763;
RN [1]
RP SEQUENCE FROM N.A.
RA Davison A.J., Akter P., Dolan A., Wright K.M., Addison C.,
RA Alcendor D.J., Hayward G.S., McGeoch D.J.;
RT "The human cytomegalovirus genome revisited.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF480884; AM00705.1; -.
DR InterPro; IPR00234; Glycoprot B.
DR InterPro; IPR000847; HTH LysR.
DR Pfam; PF00606; Glycoprotein B; 1.
DR ProDom; PD000693; Glycoprot B; 1.
DR PROSITE; PS00044; HTH LYSR FAMILY; 1.
SQ SEQUENCE 921 AA; 103982 MW; 16B63354627F55B1 CRC64;

Query Match 23.0%; Score 1047.5; DB 12; Length 921;
Best Local Similarity 28.6%; Pred. No. 1.9e-57;
Matches 273; Conservative 186; Mismatches 353; Indels 143; Gaps 25;

QY 14 LWLYQVALYSLSIAET---GVTSPTATWSTESPLTGHYGHDSHGGERGNNEKRS 69
Db 1 MWAKIWLVCVSLRTPSSGGVYCAPSSSTSQPTTPSPVPGSSSS---GRTTTRTT 56
QY 70 EQQN-----ENIYG-SPSTFPYRVCSAGVGVDFRFTDHCV----PDAS 109
Db 57 QQUESTINGTEYIFNTTLRPDDVIGVNIISKYPYRVCSMAQGTDLIRFRDNIHCTSVKPTKE 116
QY 110 DMVHSEGILLIYQNIIPFMRVRKYKVVTTSTVYNGIYSDSTINQHTFYKSIETPEWETE 169
Db 117 DL--DEGMVYVYKRNIIQAHTFKVRVQKVLTFRRSYAYIYSTLLGNTSEYVPPPPMEVH 174

```

